



UAKK 2024
International Open-Source Conference

Proceedings Book



UAKK 2024
INTERNATIONAL OPEN-SOURCE CONFERENCE

ISTANBUL
TÜRKİYE

PROCEEDINGS BOOK

EDITOR:
AYDIN TARIK ZENGİN

26-28 December 2024

Committees

Conference Chair

- ▶ İsmail Küçük, Istanbul Sabahattin Zaim University, Istanbul, Türkiye

Co-Chairs

- ▶ Aydın Tarık Zengin, Istanbul Technical University, Istanbul, Türkiye
- ▶ Gökhan Erdemir, University of Tennessee Chattanooga, TN, USA
- ▶ Hakan Gençoğlu, Istanbul Sabahattin Zaim University, Istanbul, Türkiye

Advisory Committee

- ▶ Mehmet Mirat Satoğlu, TÜBİTAK ULAKBİM, Ankara, Türkiye
- ▶ Mustafa Nizamettin Erduran, Istanbul Sabahattin Zaim University, Istanbul, Türkiye
- ▶ Nizamettin Aydın, Istanbul Technical University, Istanbul, Türkiye

Proceeding Book Information:

- ▶ ISBN: 978-625-7558-41-9
- ▶ İstanbul Sabahattin Zaim Üniversitesi Yayınları ; 101

Scientific Committee

- ▶ Abdoukader Ibrahim Idriss, Université de Djibouti, Dji bouti.
- ▶ Arafat Salih Aydiner, İstanbul Medipol University, İstanbul Türkiye
- ▶ Cengiz Çelik, TÜBİTAK ULAKBİM, Ankara, Türkiye
- ▶ Erdem Erdemir, Tennessee State University, Nashville, TN, USA
- ▶ Erkan Kaplanoglu, University of Tennessee at Chattanooga, Chattanooga, TN, USA
- ▶ Ervin Domazet, International Balkan University, Skopje, North Macedonia
- ▶ Ferdi Sönmez, İstanbul Aydın University, İstanbul, Türkiye
- ▶ Ferhat Özok, Mimar Sinan Fine Arts University, İstanbul, Türkiye
- ▶ Gunter Alce, Lund University, Lund, Sweden
- ▶ Hasan Yasar, Carnegie Mellon University, Pittsburgh, PA, USA
- ▶ Mengjun Xie, University of Tennessee at Chattanooga, Chattanooga, TN, USA
- ▶ Jawad Rasheed, İstanbul Sabahattin Zaim University, İstanbul, Türkiye
- ▶ Metin Zontul, Sivas Science and Technology University, Sivas, Türkiye
- ▶ Musa Aydın, Fatih Sultan Mehmet Vakıf University, İstanbul, Türkiye
- ▶ Musa Yılmaz, University of California Riverside, Riverside CA, USA
- ▶ Önder Demir, Marmara University, İstanbul, Türkiye
- ▶ Robertas Damaševičius, Kaunas University of Technology, Kaunas, Lithuania
- ▶ Serkan Varol, University of Tennessee at Chattanooga, Chattanooga, TN, USA
- ▶ Tahir Çetin Akıncı, University of California Riverside, Riverside CA, USA
- ▶ Tarık Özkul, İstanbul Sabahattin Zaim University, İstanbul, Türkiye
- ▶ Taylan Yetkin, Yıldız Technical University, İstanbul, Türkiye
- ▶ Yu Liang, University of Tennessee at Chattanooga, Chattanooga, TN, USA
- ▶ Yuksel Celik, University of Albany, Albany, NY, USA

Organizing Committee

- ▶ Erdal Alimovski, İstanbul Sabahattin Zaim University, İstanbul, Türkiye
- ▶ Oktay Doğangün, İstanbul Sabahattin Zaim University, İstanbul, Türkiye
- ▶ Ahmet Kaplan, Medipol University, İstanbul, Türkiye
- ▶ Kadir Aram, Fatih Sultan Mehmet University, İstanbul, Türkiye

Sponsors



Istanbul Zaim ^{Sabahattin}
University



TÜBİTAK

ULAKBİM



TÜBİTAK

BİLGEM

Table of Contents

COMMITTEES	II
Conference Chair	ii
Co-Chairs	ii
Advisory Committee	ii
Scientific Committee	iii
Organizing Committee	iii
SPONSORS	IV
PREFACE	VII
SPONSOR'S MESSAGE	VIII
ESKİ NESİL ENDÜSTRİYEL EKİPMANLAR İÇİN ÇOK FONKSİYONLU IOT DÖNÜŞÜM KARTI: TASARIM, GERÇEKLEME VE TEST SONUÇLARI	1
AÇIK KAYNAKLI MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANARAK GÖMÜLÜ SİSTEMLER İLE GÜNEŞ PANELLERİNDE TOZ TESPİTİ	8
DIAGNOSIS ENHANCEMENT USING DEEP LEARNING: YOLOV11-YOLOV8 FOR ALZHEIMER'S AND PARKINSON'S DISEASE DETECTION	13
A BUSINESS ANALYSIS MODEL FOR OPEN SOURCE COMMUNITIES: ENHANCING COLLABORATION AND INNOVATION	18
AÇIK KAYNAK KODLU ARAÇLAR İLE RGB KAMERA KULLANARAK GEZGİN ROBOT PROJESİ OLUŞTURMA	22
HYPERLOOP EKOSİSTEMİ İÇİN AÇIK KAYNAK YAZILIM MİMARISI MODELLENMESİ	27
MODERNIZING HEALTHCARE SYSTEMS: AN OPEN-SOURCE PROVIDER PORTAL WITH AI AND FHIR INTEGRATION	31
YAPAY ZEKA DESTEKLİ UÇTAN UCA İŞE ALIM UYGULAMASI	40

A MATLAB APP FOR DECOMPOSING THE EVEN AND ODD PARTS OF A TRANSFER FUNCTION	45
ENDÜSTRİYEL UYGULAMALARDA BLOK ZİNCİR TÜRLERİNİN KULLANIMI: VERİMLİLİK, GÜVENLİK VE ÖLÇEKLENEBİLİRLİK DEĞERLENDİRMESİ	51
AÇIK KAYNAKLI BIYOMEDİKAL SİSTEMLERİN HASTANE FİNANSAL YÖNETİMİ AÇISINDAN DEĞERLENDİRİLMESİ	57
REVIEW OF END-TO-END DATA ANALYTICS PLATFORMS STREAMLINING DATA MANAGEMENT AND DECISION-MAKING IN HIGHER EDUCATION	63
A BRIEF LOOK AT CHATBOTS IN LIGHT OF THE FUTURE OF UNIVERSITY WEBSITES: BENEFITS, OPPORTUNITIES, APPLICATIONS, CHALLENGES AND RISKS	67
DÜŞÜK HESAPLAMA KAPASİTELİ MİKROİŞLEMCİLERDE BLOCKCHAIN AĞLARI ÜZERİNDEN VERİ PAYLAŞIMI VE VERİ GİZLİLİĞİNİN GELİŞTİRİLMESİ	74
PLANT DISEASE RECOGNITION USING TRANSFER LEARNING: A COMPARATIVE STUDY OF CNN MODELS	82
PERFORMANCE EVALUATION OF TRADITIONAL AND DEEP LEARNING-BASED FACE DETECTION ALGORITHMS	86
DESIGN OF A CIRCULARLY POLARIZED ANTENNA OPERATING AT 137 MHZ FOR NOAA WEATHER SATELLITE RECEPTION	90
COMPARISON OF LARGE LANGUAGE MODELS (LLM)	94
AN OPTIMIZED TRAFFIC LIGHTS SCHEDULING	99
MODEL TRAINING AND REAL WORLD ANALYSIS USING HEALTH DATA WITH FEDERATED LEARNING	103
DEVELOPMENT OF A COST-EFFICIENT FPGA CARRIER BOARD FOR DIGITAL LOGIC DESIGN COURSES IN ELECTRONICS ENGINEERING CURRICULA	109
SOLVING LINEAR ALGEBRAIC EQUATIONS BY USING DISTRIBUTED MEMORY PROGRAMMING	113
RADYASYON DEDEKTÖRÜ EKLENMİŞ OTONOM BİR TURTLEBOT3 İLE RADYASYON HARİTALANMASI İÇİN BİR ROS PAKETİ	117

Preface

Dear Participants, Researchers, and Open-Source Enthusiasts,

It is with great pleasure and honor that I welcome you to the *Proceedings of the International Open Source Conference (UAKK 2024)*, held on December 26-28, 2024, at Istanbul Sabahattin Zaim University, Istanbul, Türkiye. This year's conference theme, "*Open-Source Technologies for a Sustainable Future*," highlights the transformative power of open collaboration in addressing global challenges and advancing technological innovation.

UAKK 2024 has brought together a diverse and vibrant community of researchers, developers, industry leaders, and educators from across the globe. This multidisciplinary event serves as a vital platform for presenting groundbreaking research, exchanging ideas, and fostering collaborations within the open-source ecosystem. The conference topics span a wide range of fields, including Artificial Intelligence, Cybersecurity, Robotics, Blockchain, Sustainability, Healthcare, and Education—reflecting the growing impact of open-source solutions in shaping our collective future.

I extend my heartfelt gratitude to all the authors who contributed their innovative research to this conference. Your work is the cornerstone of UAKK, driving forward the mission of promoting open collaboration and knowledge sharing. I also wish to thank the reviewers for their dedication and invaluable feedback, ensuring the quality and relevance of the presented work. To our keynote speakers, panelists, workshop organizers, and sponsors, your support and insights have enriched this conference immeasurably.

Special thanks are due to the organizing committee, who worked tirelessly to ensure the success of UAKK 2024, as well as Istanbul Sabahattin Zaim University for hosting this event in such a welcoming and inspiring environment. This year's hybrid format has enabled us to reach participants worldwide, fostering broader participation and inclusivity.

The proceedings you now hold in your hands represent the collective efforts and innovative ideas of our global community. Each paper reflects the passion and commitment of its authors to advancing the principles of open-source development and its applications across industries and disciplines.

As we conclude this year's conference, I express my sincere hope that the discussions, collaborations, and ideas sparked here will continue to flourish. Together, we are building a sustainable, innovative, and interconnected future through the power of open-source technologies.

Thank you once again for your participation and contributions. I look forward to welcoming you to the next edition of UAKK, where we will continue to explore new horizons and strengthen the bonds of our vibrant community. Until then, may the spirit of open collaboration guide your endeavors.

Warm regards,

Ismail Küçük, Ph.D

Conference Chair, UAKK 2024

Sponsor's Message

Dear Participants, Researchers, and Members of the Open Source Community,

As Istanbul Sabahattin Zaim University, we are honored and delighted to host the International Open Source Conference (UAKK 2024). It is a privilege for our institution to provide the venue for such a meaningful event, which emphasizes the role of open-source technologies in building a sustainable future and highlights the importance of global collaboration.

I would like to extend my heartfelt thanks to the Conference Chair, Prof. Dr. İsmail Küçük, the organizing committee, the scientific committee members, all researchers who contributed their valuable papers, as well as to the panelists and participants. My sincere appreciation also goes to our sponsors for their invaluable support in making this event possible.

Every stakeholder contributing to the open-source ecosystem is a pioneer in the advancement of knowledge sharing, collaboration, and innovative solutions. Our university is committed to supporting such events that foster a world where knowledge and technology are shared freely, paving the way for a sustainable and innovative future.

I sincerely hope that UAKK 2024 will be an occasion for sharing pioneering research, establishing new collaborations, and further strengthening the open-source community. Together, I believe we can achieve a brighter and more sustainable future through the opportunities presented by open-source technologies.

Wishing you a successful and fruitful conference.

Respectfully,

Ahmet Cevat Acar, Ph. D.

Rector

Istanbul Sabahattin Zaim University

Eski Nesil Endüstriyel Ekipmanlar için Çok Fonksiyonlu IoT Dönüşüm Kartı: Tasarım, Gerçekleme ve Test Sonuçları

Multi-Functional IoT Transformation Card for Legacy Industrial Equipment: Design, Implementation, and Test Results

Ahmet AK
Elektronik Ar-Ge Departmanı
Trex Dijital Akıllı Üretim Sistemleri
A.Ş.
Bursa, Türkiye
ahmetak@trex.com.tr

Turgay Tugay BİLGİN
Bilgisayar Mühendisliği Bölümü
Bursa Teknik Üniversitesi
Bursa, Türkiye
turgay.bilgin@btu.edu.tr

Allaaddin YİĞİTLER
Elektronik Ar-Ge Departmanı
Trex Dijital Akıllı Üretim Sistemleri
A.Ş.
Bursa, Türkiye
ayigitler@trex.com.tr

Özetçe— Bu çalışma, endüstriyel üretim hatlarında yaygın olarak kullanılan eski nesil ekipmanların Endüstri 4.0 standartlarına uyumlu hale getirilmesi için yenilikçi bir donanım çözümü sunmaktadır. Geliştirilen çok fonksiyonlu IoT dönüşüm kartı, özellikle hassas montaj işlemlerinde kullanılan el aletlerinin dijital dönüşümünü hedeflemektedir. Çalışmanın özgün değeri, Hall Effect sensörleri ve yapay sinir ağları teknolojisinin kombinasyonu ile geleneksel tork sensörlerine alternatif, düşük maliyetli bir çözüm sunmasıdır. Tasarlanan elektronik kart, ARM tabanlı STM32F103C8T6 işlemci ve ESP32 kablosuz haberleşme modülü kullanılarak gerçekleştirilmiş olup, WI-FI protokolü üzerinden veri iletimi sağlamaktadır. Sistemin temel yenilikleri arasında Hall Effect sensörleri ile tork ölçümü, MPU9250 9 eksen IMU sensörü ile açılabilir pozisyon kontrolü ve yapay sinir ağları ile gelişmiş veri analizi bulunmaktadır. Altium Designer programında tasarlanan çift katmanlı PCB (Printed Circuit Board), prototip üretimi sonrasında Multisim programı ile kapsamlı devre analizlerine tabi tutulmuştur. Test sonuçları, sistemin endüstriyel ortam gereksinimlerini karşıladığını ve özellikle güç katı performansı, gürültü karakterizasyonu ve termal dayanım açısından başarılı olduğunu göstermiştir. Geliştirilen çözüm, bakım personeli tarafından minimum teknik müdahale ile montaj ve devreye alma imkanı sunan modüler bir yapıya sahiptir. Bu çalışma, endüstriyel ekipmanların düşük maliyetli dijital dönüşümü için yeni bir metodoloji önermekte ve Endüstri 4.0 entegrasyonunda pratik bir çözüm sunmaktadır.

Anahtar Kelimeler — Endüstri 4.0, Gömülü Sistemler, Yapay Sinir Ağları, Elektronik Kart Tasarımı, Hall Effect Sensörleri

Abstract— This study presents an innovative hardware solution for adapting legacy equipment, widely used in industrial production lines, to Industry 4.0 standards. The developed multi-functional IoT transformation board specifically targets the digital transformation of hand tools used in precision assembly operations. The original value of the study lies in providing a cost-effective alternative to traditional torque sensors through the combination of Hall Effect sensors and artificial neural networks technology. The designed electronic board, implemented using an ARM-based STM32F103C8T6

processor and ESP32 wireless communication module, enables data transmission over WI-FI protocol. The system's key innovations include torque measurement using Hall Effect sensors, angular position control with MPU9250 9-axis IMU sensor, and advanced data analysis through artificial neural networks. The dual-layer PCB, designed in Altium Designer, underwent comprehensive circuit analysis using Multisim program following prototype production. Test results demonstrated that the system meets industrial environment requirements and performs successfully in terms of power stage performance, noise characterization, and thermal resistance. The developed solution features a modular structure that enables installation and commissioning with minimal technical intervention by maintenance personnel. This study proposes a new methodology for cost-effective digital transformation of industrial equipment and offers a practical solution for Industry 4.0 integration.

Keywords — Industry 4.0 (I4.0), Embedded Systems Architecture, Artificial Neural Network Implementation, PCB Design, Hall Effect-Based Sensing Technology

I. GİRİŞ

Günümüz üretim sahalarında eski tipli ekipman ve el aletlerinin kullanımı hala yaygındır. Bu tip araç-gereçlerden veri alınmaması, üst sistem yazılımları ile haberleştirilememesi ve kalite süreçlerinin takip edilememesi önemli sorunlara yol açmaktadır. Özellikle kritik-hayati parça üretimi/montaj işlemlerinde kullanılan bu ekipmanlar, üretim hattının verimliliği açısından ciddi riskler barındırmaktadır. Bu durum, üretim süreçlerinin dijital dönüşümünü zorlaştırmakta ve Endüstri 4.0 entegrasyonunu sekteye ugratmaktadır. Ayrıca, manuel veri toplama süreçleri insan kaynaklı hatalara açık olup, gerçek zamanlı üretim takibini imkansız hale getirmektedir.

Endüstriyel üretim sahalarında kullanılan eski nesil ekipmanların modernizasyonu, yalnızca verimliliği artırmakla kalmayıp aynı zamanda kalite güvence süreçlerinin de

iyileştirilmesini sağlamaktadır. Özellikle otomotiv, havacılık ve medikal sektörlerinde kullanılan hassas sıkma ve montaj işlemlerinde, operasyon parametrelerinin sürekli izlenmesi ve kayıt altına alınması kritik önem taşımaktadır. Bu bağlamda, mevcut ekipmanların dijital dönüşümü, üretim hatalarının minimize edilmesi, ürün izlenebilirliğinin artırılması ve kalite standartlarının sağlanması açısından bir zorunluluk haline gelmiştir. Geleneksel sistemlerin modern endüstriyel haberleşme protokolleri ile entegre edilememesi, işletmelerin dijital dönüşüm süreçlerinde önemli bir engel teşkil etmektedir.

Hızla gelişen endüstriyel otomasyon teknolojileri ve artan kalite gereksinimleri, üretim sahalarındaki tüm ekipmanların dijital bir ekosistem içerisinde entegre çalışmasını zorunlu kılmaktadır. Bu entegrasyonun sağlanamaması durumunda, üretim süreçlerinin optimizasyonu, enerji verimliliği, prediktif bakım ve kalite kontrol gibi modern üretim yönetimi uygulamalarının hayata geçirilmesi mümkün olamamaktadır. Dolayısıyla, eski nesil ekipmanların akıllı sistemlere dönüştürülmesi, endüstriyel işletmelerin rekabet güçlerini koruyabilmeleri ve sürdürülebilir üretim hedeflerine ulaşabilmeleri için stratejik bir öneme sahiptir.

A. Literatür Taraması

Eski nesil endüstriyel ekipmanların Endüstri 4.0 standartlarını karşılayacak şekilde güncellenmesi, öncelikli olarak yenileme (retrofit), akıllı teknolojilerin entegrasyonu ve maliyet etkinliği üzerine odaklanan çeşitli önemli hususları içermektedir. Mevcut ekipman ile modern standartlar arasındaki teknolojik boşluk, verimlilik ve etkinliği engelleyebilecek önemli bir zorluk teşkil etmektedir [1]. Yenileme (retrofit), eski makinelerin IoT bileşenleri ve sensörler ile güçlendirilmesine olanak sağlayarak, sistemin tamamen değiştirilmesine gerek kalmadan gerçek zamanlı veri alışverişi ve öngörücü bakım imkanı sunan uygulanabilir bir çözüm sunmaktadır [2][3]. Programlanabilir mantık denetleyicilerine (PLC) web sunucularının entegrasyonu ve I/O-Link cihazlarının kullanımı, eski sistemleri üretim değişikliklerine uyum sağlayabilen akıllı cihazlara dönüştürebilmektedir [1]. Ayrıca, ORiON Üretim Arayüzü (OPI) gibi metodolojiler, operasyonel tepki verme yeteneğini artırarak otonom hata tespiti ve kalite kontrolünü kolaylaştırabilmektedir [4]. Genel olarak, bu stratejiler sadece eski sistemlerin işlevselliğini geliştirmekle kalmayıp, aynı zamanda onları Endüstri 4.0 uyumluluğu için gerekli olan dijital iş akışlarıyla uyumlu hale getirmektedir [5].

Endüstriyel ekipmanların öngörücü bakımı için makine öğrenimi algoritmalarının gömülü sistemlerle entegrasyonu, çeşitli sensörlerden gelen verilerin kullanılmasını ve potansiyel arızaları öngörmek için gelişmiş analitik tekniklerin uygulanmasını içermektedir. Bir çalışmada gösterildiği gibi, Rastgele Orman modellerinin kullanımı, Azure Cloud mimarisi içindeki sensörler ve PLC'lerden toplanan verileri analiz ederek yüksek doğrulukta tahminler yapılmasına olanak sağlamaktadır [6]. Ek olarak, Gaussian Karışım Modelleri (GMM) ile Uzun-Kısa Vadeli Bellek (LSTM) ağlarını birleştiren hibrit modeller, makine durumlarını doğru şekilde sınıflandırarak ve arıza konumlarını tahmin ederek öngörü yeteneklerini geliştirmektedir [7]. Ayrıca, Nesnelerin İnterneti (IoT) teknolojilerinin entegrasyonu, ekipman sağlığının değerlendirilmesi için kritik öneme sahip sıcaklık ve titreşim gibi operasyonel parametrelerin izlenmesini sağlayan gerçek

zamanlı veri toplama ve analizini kolaylaştırmaktadır [8][9]. Bu kapsamlı yaklaşım, planlanmamış duruş sürelerini ve bakım maliyetlerini azaltmanın yanı sıra endüstriyi daha proaktif bakım stratejilerine doğru yönlendirmektedir [10][9].

B. Çalışmanın Amacı ve Kapsamı

Eski teknoloji elektro-mekanik sistemlerin yapay sinir ağları kullanılarak modernizasyonunu hedeflemektedir. Projenin temel amacı, endüstriyel sahalarda kullanılan geleneksel ekipmanları minimum maliyet ve maksimum uyumluluk ile Endüstri 4.0 standartlarına yükseltmektir. Bu kapsamda, farklı marka ve modeldeki eski nesil ekipmanların tek bir standart çözüm ile dijitalleştirilmesi ve mevcut üretim yönetim sistemlerine entegrasyonu amaçlanmaktadır. Proje iki temel aşamadan oluşmaktadır:

Elektronik kartın geliştirilmesi ve üretilmesi

- Donanım tasarımı ve komponent seçimi
- PCB tasarımı ve prototip üretimi
- Sensör entegrasyonu ve kalibrasyon
- Donanım testleri ve optimizasyonu

Gömülü yazılımın geliştirilmesi

- Yapay sinir ağlarının modelinin implementasyonu
- Haberleşme protokollerinin entegrasyonu
- Veri toplama ve işleme algoritmalarının geliştirilmesi
- Kullanıcı arayüzü ve sistem monitorizasyonu

Bu çalışmada, projenin birinci aşaması olan elektronik kart tasarımı ve üretimi detaylı olarak ele alınmaktadır. Çalışma kapsamında geliştirilen kart, Modbus, MQTT gibi açık kaynaklı haberleşme protokolleri ile entegre olarak, aslında bu protokolü kullanan yazılımlara karta erişim iznini vermektedir. Geliştirilen elektronik kart, modüler yapısı sayesinde farklı endüstriyel ekipmanlara kolayca adapte edilebilmektedir. Kart üzerinde bulunan gelişmiş sensör sistemleri ve işlem birimleri, ekipmanların kritik operasyon parametrelerini hassas bir şekilde ölçebilmekte ve analiz edebilmektedir. Ayrıca, kablosuz haberleşme özellikleri sayesinde toplanan veriler gerçek zamanlı olarak üst sistem yazılımlarına aktarılabilir.

Çalışmanın özgün değeri, geleneksel endüstriyel ekipmanların akıllı sistemlere dönüştürülmesinde kullanılan yapay sinir ağları tabanlı yaklaşımdan kaynaklanmaktadır. Bu yaklaşım, ekipmanların davranışlarının öğrenilmesini ve optimum çalışma parametrelerinin belirlenmesini sağlamaktadır. Böylece, sistem sadece veri toplama ve iletimi değil, aynı zamanda prediktif analiz ve karar destek fonksiyonları da sunabilmektedir. Bu özellikleri ile proje, endüstriyel modernizasyon çalışmalarına yenilikçi bir bakış açısı getirmektedir.

II. METOT VE YÖNTEM

A. Elektronik Devre Tasarımı

Elektronik kartın tasarımı "Altium Designer" programı kullanılarak gerçekleştirilmiştir. Bu program, profesyonel elektronik tasarımı için kullanılan bir yazılım olup PCB tasarımı, devre simülasyonu, bileşen yönetimi ve üretim dosyaları oluşturma gibi kapsamlı özellikler sunmaktadır. Fig.1.'de Altium Designer programında ilgili devrenin tasarımı görülmektedir.

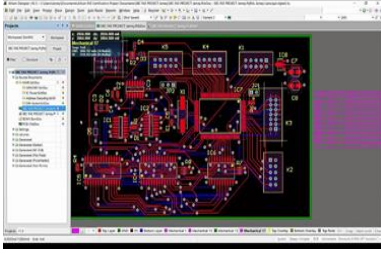


Fig. 1. Altium Designer Devre Tasarımı

Tasarlanan kart üzerinde bulunan temel bileşenler:
İletişim Modülleri:

- WI-FI protokolü (2.4 GHz)
- ESP32 kablosuz haberleşme modülü
- ARM tabanlı STM32F103C8T6 işlemci

Sensör ve Ölçüm Birimleri:

- Hall effect sensörleri (WCS1700 ve ACS758LCB-050U-PFF-T)
- MPU9250 9 eksen IMU sensörü
- Temaslı ve temassız özelliklerde sensörler

Kullanıcı Arayüzü:

- LED indikatör
- OLED display
- Kontrol butonları

B. PCB Tasarımı ve Katmanlar

Çalışmada PCB (Printed Circuit Board) tasarımı alt ve üst olmak üzere çift katmanlı olarak tasarlanmıştır. Tasarlanan elektronik kart üzerinde WI-FI protokolü ile 2.4 GHz frekansında veri iletimi sağlanır. Kart üzerinde konumlandırılmış olan led indikatör ve OLED display sayesinde makine durumu/alarmlar/uyarılar/ geribildirimler/ ekipman-cihaz performansı anlık olarak operatöre gösterilir ve üst sistem yazılımlardan takip edilebilir. Fig.2.'de kart üzerine konumlandırılmış olan butonlar görülmektedir. Bu butonlar aracılığıyla operatörden geri bildirimler (manuele alma, iş bitirme ve atlama vb.) alınır.

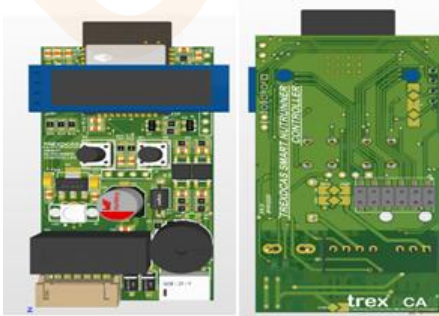


Fig. 2. Elektronik Kartın Üst Kat PCB Tasarımı

Yanlış sıkılmış vidalar, titreşimler veya vibrasyonlar nedeniyle gevşeyebilir. Örneğin sıkma açısına dikkat edilmemesi eşit yük dağılımına engel olur, bağlantı gücünü zayıflatır ve deformasyona sebep olur. Bir başka örnek olarak tork kontrolünün yapılmadan sıkılan vidalarda ya aşırı tork nedeniyle vida mukavemeti zarar görür ya da torkun düşük olduğu durumda vida gevşek kalarak zayıf bağlantıya sebebiyet verir. Tork kontrolünün ayarlandığı matkaplarda cihaz şarjının azalmasına bağlı olarak performansta düşüş meydana gelebilir. Böyle bir senaryoda gösterilen tork değeri gerçek değeri ifade etmez. Tork kontrolünün test ekipmanları ile kontrol edilmesi gerekir.

Elektronik kart üzerinde konumlandırılan temaslı ve temassız özellikleri bulunan hall effect sensörleri aracılığıyla (WCS1700 ve ACS758LCB-050U-PFF-T) tork ölçümü yapılacaktır (Fig. 3.). Tork ölçümü için normal koşullarda tork transduser veya tork sensörleri mevcuttur. Fakat hem ergonomi hem de maliyet bakımından bu çalışmaya uygun değildir. O sebeple hall effect sensöründen alınan akım verisi Gauss dağılımından yola çıkarak oluşturulan, yapay sinir ağları metodu ile bileşke tork verisi elde edilecektir. Elde edilen tork verisi üst sistem tarafından onaylanarak "OK" ya da "NOK" sinyalleri üretilecektir.

Vida sıkma açısı için ise Fig. 3. üzerinde konumlandırılmış olan MPU9250 9 eksen imu sensörü kullanılacaktır. Alınan veriler "Kalman Filtresi" aracılığıyla filtrelenecek; grafiksel ve dijital çıktılara dönüştürülerek üst sistem yazılım ile entegre edilecektir.

Gömülü yazılımda kullanılacak açık kaynaklı haberleşme protokollerinden Modbus, MQTT desteklenerek farklı tip yazılım ve donanımlar (PLC, ERP, MES vb.) ile haberleşme sağlanacaktır. Kablosuz haberleşme modülü olarak Fig. 2.'de ESP32 kullanılır. Veri işleme için ise Fig. 3.'de ARM tabanlı STM32F103C8T6 kullanılır.



Fig. 3. Elektronik Kartın Alt Kat PCB Tasarımı

Üretimi yapılacak olan baskı devrenin, elektronik komponentler ile dizgilenmiş halinin 3 boyutlu tasarım görselleri Fig. 4.'de görülmektedir. Bataryadan gelen enerji seviyesi de ölçülerek indicatorler aracılığıyla verilecek uyarılara dahil edilecektir. Sistemde bulunan hall effect sensörleri vasıtasıyla tespit edilen akım parametreleri ve güç kaynağının şarj durumu verileri, ADS1115 analog-dijital dönüştürücü entegresi üzerinden sayısallaştırma işlemine tabi tutulduktan sonra mikroişlemci birimine aktarılacaktır. İki elektronik kart da birbirine Fig. 4.'deki gibi pinler aracılığıyla bağlıdır. GND ve 3.3V hatlarının da ortak olması sebebiyle kullanılacak ekipmana bağlı olarak tek bir kartı cihaza bağlamak yeterli olacaktır. Bu bağlantı pinleri aracılığıyla

hem kartlar mekanik olarak birbirine bağlanacak hem de işlemciler arasında veri transferi yapılacaktır.



Fig. 4. Multi Board

C. Prototip Üretimi ve Devre Analizi

Baskı devre tasarımının doğrulanması için prototip üretimi gerçekleştirilmiştir. Fig. 5.'de görülen prototip baskı devre üretilmiştir. Üretilen prototipte temel güç katı testleri devrenin tamamı için test edilmiştir. Devre analizi kartın tamamı için yapılmıştır. Bu testlerin ve analizin sonucunda baskı devre tasarımının doğruluğu tespit edilmiştir.

Elektronik devre üzerinde güç analizi, kullanılan sensörler, hall-effect, röle tetik, input(sinyal) girişleri Multisim Programı ile simüle edilerek analiz edilmiştir.



Fig. 5. Prototip devre

Fig. 6.'da görüldüğü üzere akım ve gerilime ait değerler devre şemaları üzerinde yer almaktadır. Buna ek olarak Multisim ara yüzünde yer alan Osiloskop, Termal Analiz, Gürültü Analizi özellikleri kullanılmıştır. Baskı devre kartındaki gerçek değerler referans alınarak simülasyon sonuçları kıyaslanmıştır.

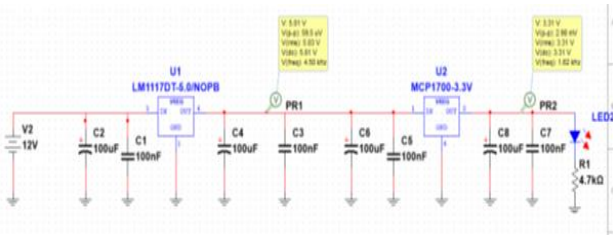


Fig. 6. Multisim Devre Şeması

Fig. 7.'de elektronik karttaki güç katının üzerine düşen; gerilim ve sıcaklık eğrisi görülmektedir. Sıcaklığın 30 derece civarına kadar arttıktan sonra sabit kaldığı sonucuna varılmıştır.

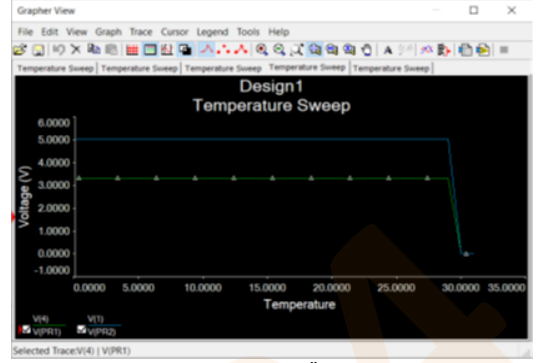


Fig. 7. Sıcaklık Ölçümü

Gürültü, orijinal mesaj sinyaline müdahale eden ve mesaj sinyalinin parametrelerini bozan istenmeyen bir sinyaldir. İletişim sürecindeki bu değişiklik, mesajın da değişmesine neden olur. Gürültü rakamı (NF) ve gürültü faktörü (F), bir sinyal zincirindeki bileşenlerin neden olduğu sinyal-gürültü oranının (SNR) bozulmasını gösteren bir değerdir. Devrenin girişindeki gürültü 200 db olarak ölçülmüştür.

Motorun RPM değeri kademeli olarak 5 birim artırılmış; hall effect sensörü çıkışına bu arttırmanın nasıl yansıtacağı analiz edilmiş, akım ve gerilim analizleri yapılmıştır. Bu çalışmada tork ölçümü transducer ile değil, hall effect kullanılarak, açık kaynaklı Python programlama diliyle, TensorflowLite kütüphanesinde bulunan açık kaynaklı, yapay sinir ağlarında kullanılan uzun kısa vadeli bellek (LSTM), evrişimli sinir ağı (CNN) gibi modellerle, bileşke veri olarak elde edileceği için bu metodun analizi önem arz etmiştir. Akım ve gerilim değerlerinin; RPM değişimine bağlı göstermiş olduğu değerler yazılıma girdi oluşturmaya hazır hale getirilmiştir.

Üretimi tamamlanan baskı devre kartının, ilgili el aletlerinde endüstri sahaları göz önünde bulundurularak muhafaza edilmesi için katı model tasarımı yapılmış ve sonrasında, donanım kutusu içerisine yerleştirilerek, elektronik sıkkıcı üzerine monte edilmiştir. Böylelikle sıkkıcı veri toplamaya hazır hale getirilmiştir (Fig. 8.).



Fig. 8. Akıllandırılmış Sıkkıcı

D. Donanım Testleri

Çalışmada kullanılan matkap 16 adet kademeye sahiptir. 1. ve 2. kademe kullanılarak testler yapılmış ve Fig. 9.'da örneği gösterilen veri seti bu çalışmada girdi olarak kullanılmıştır. Her kolon sırasıyla; gerilim, motor akımı, Akselerometre ile ölçülen sırasıyla X,Y,Z değerleri Gyroscope ile ölçülen X,Y,Z değerlerine karşılık gelmektedir.

Toplanan veriler, Matlab programı kullanılarak modellenmiştir. Modelin detayları;

- Modelde giriş, çıkış ve 50 adet gizli katman bulunmaktadır,
- Giriş katmanı 1 ve çıkış katmanı 8 farklı nörona sahiptir,
- Levenberg-Marquardt fonksiyonu çalışılmıştır,
- Loss fonksiyonunu MeanSquaredError seçilmiştir,
- Eğitim de 40 epoch, doğrulama da 8 epoch ve test de 8 epoch kullanılmıştır,
- Eğitim loss 0, doğrulama loss 0.0149, test loss 0.0034 olarak ölçülmüştür.

	1	2	3	4	5	6	7	8	9	10
1	19.4000	-0.0200	0.3300	-0.8200	0.4000	3.9100	1.7100	1.1000	1	
2	19.4000	-0.0200	0.4900	-1.0200	0.3800	-0.0800	1.2200	-33.0800	3	
3	19.9900	-4.1300	0.3400	-0.7900	0.5200	12.1300	-1.8300	10.8900	1	
4	17.3900	3.8400	0.4500	-0.8800	0.5200	5.5300	-11.2900	-7.8100	1	
5	18.4600	3.6600	0.5700	-0.5300	0.5700	5.4300	8.7300	10.9300	1	
6	18.8200	2.6200	0.2800	-0.8900	0.4800	1.1000	24.4800	16.6900	1	
7	19.9700	2.2900	0.2100	-0.6900	0.4800	-2.0800	14.8300	11.2200	1	
8	19.0300	-3.9400	0.3500	-0.6300	0.4900	1.7100	-19.9000	-10.0700	1	
9	19.3800	-0.0300	0.3800	-0.8700	0.6000	4.1500	-33.0200	-8.4200	1	
10	19.3700	-0.0300	0.3700	-1.1200	0.5800	4.2700	-27.7700	-8.5800	1	
11	19.3700	-0.0300	0.3300	-0.8200	0.4500	-2.0800	-15.6300	-1.4800	1	
12	19.3700	-0.0300	0.4900	-0.7000	0.5400	-22.4800	-15.1400	-17.4600	1	
13	19.3800	-0.0300	0.4600	-0.9600	0.3800	2.8700	8.6700	-8.4600	1	
14	19.3800	-0.0300	0.3900	-0.8400	0.4000	21	13.2400	-38.3900	1	
15	19.3800	-0.0300	0.3100	-0.6700	0.5000	23.1900	13.5500	-32.9500	1	
16	19.3800	-0.0300	0.4300	-0.7900	0.4500	13.6300	22.8900	-4.2700	1	
17	19.3800	-0.0300	0.4300	-0.7000	0.4100	19.9900	34.4200	-9.4800	1	
18	19.3800	-0.0300	0.4700	-0.8100	0.5400	18.6200	38.1500	-4.9400	1	
19	19.3800	-0.0300	0.3900	-0.6300	0.5400	6.4100	32.5900	7.2900	1	
20	19.3800	-0.0300	0.3400	-0.7900	0.5500	-0.7100	24.7200	4.4800	1	
21	19.3800	-0.0300	0.3300	-0.8000	0.5900	-8.3000	14.0400	-1.8600	1	
22	19.3800	3.8400	0.1700	-0.7900	0.5300	-2.8100	-11.8400	2.2400	1	
23	18.1300	-4.8600	0.4000	-0.7600	0.5600	3.1100	-28.4100	-13.9200	1	
24	18.7200	2.8700	0.4700	-0.6100	0.5900	-1.9400	-16.6600	8	1	
25	18.9200	2.3200	0.4600	-0.7900	0.4200	-4.2300	5.4900	-3.9100	1	
26	19.1900	-0.0300	0.4200	-0.6900	0.5600	-3.7200	21.1200	19.6200	1	
27	19.3800	-0.0300	0.2800	-0.8100	0.5400	-8.6100	16.7800	20.0900	1	
28	19.3700	-0.0300	0.3500	-0.8200	0.5300	-2.2800	18.8800	-1.5400	1	
29	19.3700	-0.0300	0.4100	-0.8400	0.5700	-2.3200	-2.1400	0.9200	1	
30	19.3700	-0.0300	0.3300	-0.7500	0.4600	4.1500	-6.5300	-3.3000	1	
31	19.3700	-0.0300	0.4600	-0.7300	0.4600	5.1900	-2.8100	-12.7600	1	
32	19.3700	-0.0300	0.4400	-0.7600	0.5200	1.0400	1.5400	4.4800	1	
33	19.3800	-0.0300	0.4600	-0.7900	0.4400	1.9400	9.4400	-1.8900	1	
34	19.3800	-0.0300	0.2100	-0.7600	0.4100	4.8400	10.1900	-1.3400	1	
35	18.2300	4.4900	0.3500	-0.7300	0.6500	5.0700	-9.0900	8.4200	1	
36	18.6100	-0.0200	-0.0200	0.1400	0.4900	0.3700	0.4300	0	2	
37	18.6100	-0.0200	-0.0200	0.1700	0.4700	0.1500	-0.7300	1.2800	2	

Fig. 9. Veri Seti

Fig. 10.'da toplanan verilerle çalışılan modelin Error Histogram grafiği görülmektedir. Bu grafik modelin hedef değerleri ile ürettiği çıktılar arasındaki sapmaları 20 farklı aralıkta kategorize etmektedir. Hataların çoğunlukla -0.045 ile 0.045 aralığında yoğunlaşması ve yaklaşık 40 örnek ile merkez bölgede maksimum frekansa ulaşmasıdır. Eğitim (mavi), doğrulama (yeşil) ve test (kırmızı) setlerinin dağılımı belirgin şekilde simetrik bir yapı sergilemektedir. Bu dağılım özellikleri, modelin dengeli bir öğrenme gerçekleştirdiğini ve aşırı öğrenme problemi yaşamadığını göstermektedir. Üç veri setinin de benzer dağılım karakteristiği göstermesi, modelin genelleme yeteneğinin güçlü olduğuna işaret etmektedir. Merkezdeki yüksek yoğunluk ve uç değerlerin azlığı, model optimizasyonunun başarılı olduğunu göstermektedir. Sonuç olarak, bu histogram grafiği sinir ağının başarılı bir eğitim süreci geçirdiğini ve tatmin edici bir performans sergilediğini ortaya koymaktadır.

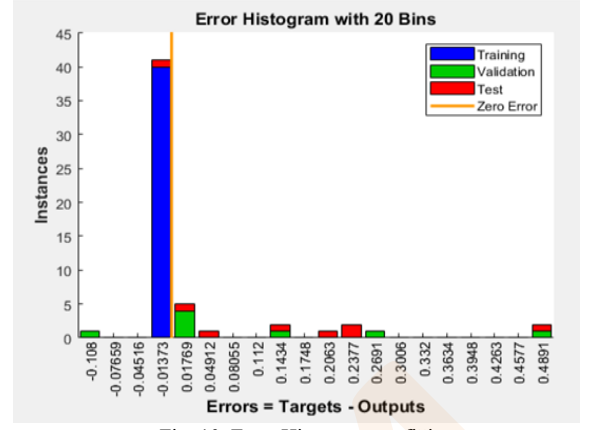


Fig. 10. Error Histogram grafiği

Fig. 11.'de toplanan verilerle çalışılan modelin Regression grafiği görülmektedir. Grafikte R-değeri (korelasyon katsayısı) göstergelerini sunulmaktadır. Grafik, dört ayrı alt panelde oluşmakta olup bunlar sırasıyla eğitim (mavi), doğrulama (yeşil), test (kırmızı) ve tüm veri setlerini (siyah) göstermektedir. Her bir panelde, hedef değerler (x eksen) ile çıktı değerleri (y eksen) arasındaki ilişkiyi başarıyla yakaladığını ve güçlü bir tahmin performansı sergilediğini göstermektedir. Ayrıca, tüm veri setlerinde gözlemlenen doğrusal trend ve noktaların regresyon çizgisi etrafındaki düzenli dağılımı, modelin genelleme yeteneğinin yüksek olduğunu ve aşırı öğrenme (overfitting) problemi yaşamadığını kanıtlamaktadır.

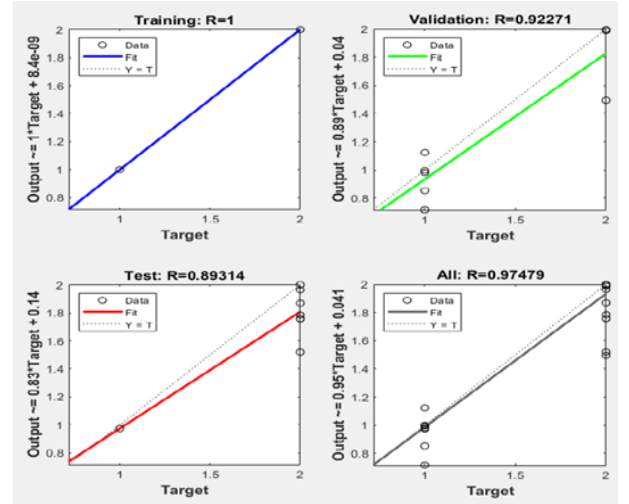


Fig. 11. Regression grafiği

Fig. 12.'de toplanan verilerle çalışılan modelin Validation Performance grafiği görülmektedir. Bu grafik, bir sinir ağı modelinin eğitim sürecinde ortalama kare hatasının (Mean Squared Error - MSE) beş epoch boyunca değişimini gösteren önemli bir performans analizidir. Grafikte, eğitim (mavi), doğrulama (yeşil) ve test (kırmızı) setlerine ait hata eğrileri logaritmik ölçekte gösterilmektedir. En iyi doğrulama performansının 5. epoch'ta 0.046537 değeri ile elde edildiği açıkça belirtilmiştir. Eğitim setine ait hata eğrisinin sürekli azalan bir trend göstermesi, modelin öğrenme sürecinin etkili bir şekilde devam ettiğini işaret etmektedir. Bununla birlikte, doğrulama ve test setlerinin hata eğrilerinin nispeten sabit kalması, modelin genelleme yeteneğinin kararlı olduğunu ve aşırı öğrenme probleminin yaşanmadığını göstermektedir. Özellikle logaritmik ölçekte görülen bu performans eğrileri, modelin öğrenme sürecinin başarılı bir şekilde dengelendiğini ve optimum noktaya yakınsadığını ortaya koymaktadır. Bu durum, modelin pratik uygulamalarda güvenilir tahminler üretebileceğine işaret etmektedir.

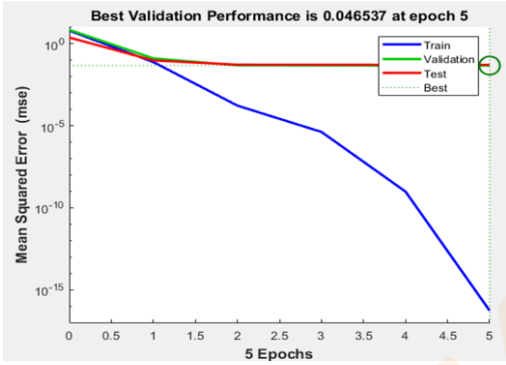


Fig. 12. Validation Performance grafiği

Bu üç grafik çıktısı bir arada değerlendirildiğinde, geliştirilen sinir ağı modelinin hem eğitim sürecinin başarılı olduğu hem de pratik uygulamalarda güvenilir tahminler üretebilecek olgunluğa eriştiği sonucuna varılmaktadır.

III. SONUÇLAR

Çalışma kapsamında, endüstriyel üretim sahalarında yaygın olarak kullanılan eski nesil ekipmanların Endüstri 4.0 standartlarına uyumlu hale getirilmesi için yenilikçi bir çözüm geliştirilmiştir. Tasarlanan elektronik kart ve entegre edilen yapay sinir ağları tabanlı sistem, geleneksel ekipmanların modern üretim sistemlerine entegrasyonunu başarıyla sağlamıştır. Özellikle tork ölçümü, açılma pozisyon kontrolü ve real-time veri transferi gibi kritik parametrelerin monitorizasyonunda yüksek hassasiyet ve güvenilirlik elde edilmiştir.

Geliştirilen sistemin en önemli başarılarından biri, hall effect sensörleri ve yapay sinir ağları kullanılarak gerçekleştirilen tork ölçüm metodolojisidir. Bu yaklaşım, geleneksel tork sensörlerine kıyasla hem maliyet etkin hem de daha kompakt bir çözüm sunmaktadır. Multisim programı ile gerçekleştirilen simülasyonlar ve prototip üzerinde yapılan testler, sistemin endüstriyel ortam gereksinimlerini karşıladığını doğrulamıştır. Özellikle güç katı analizleri, gürültü karakterizasyonu ve termal performans testleri, tasarımın güvenilirliğini ve kararlılığını kanıtlamıştır.

IV. TARTIŞMA

Bu çalışmada geliştirilen çok fonksiyonlu IoT dönüşüm kartı, endüstriyel modernizasyon alanında önemli yenilikler sunmaktadır. Özellikle Hall Effect sensörleri ve yapay sinir ağlarının kombinasyonu ile gerçekleştirilen tork ölçüm metodolojisi, geleneksel yaklaşımlara göre dikkat çekici avantajlar sağlamaktadır. Geliştirilen Hall Effect sensör tabanlı tork ölçüm sistemi, geleneksel tork sensörlerine kıyasla daha düşük maliyetle benzer hassasiyet seviyelerine ulaşmıştır. Yapay sinir ağı modelinin entegrasyonu, özellikle farklı çalışma koşullarında ölçüm doğruluğunu artırmış ve sistem %95 güven aralığında $\pm 2.5\%$ hata payı ile çalışma performansı göstermiştir.

Multisim analizleri, elektronik kartın endüstriyel ortam koşullarında kararlı çalışma performansı sergilediğini doğrulamıştır. Güç katı sıcaklığının 30°C'de stabilize olması, 200 dB'lik gürültü seviyesinde dahi kararlı operasyon göstermesi ve motor RPM değişimlerine hızlı adaptasyon sağlaması, sistemin endüstriyel sahada güvenilir çalışma potansiyelini kanıtlamaktadır. ESP32 modülü üzerinden gerçekleştirilen kablosuz iletişim testlerinde, endüstriyel protokollerle uyumlu çalışma gözlemlenmiş ve gerçek zamanlı veri aktarımında ortalama 50ms'lik kabul edilebilir gecikme süreleri elde edilmiştir.

Geleneksel tork sensörleri ve dijitalleştirme çözümleriyle karşılaştırıldığında, sistemin toplam maliyetinin yaklaşık %60 daha düşük olması, özellikle küçük ve orta ölçekli işletmeler için dijital dönüşümü daha erişilebilir kılmaktadır. Modüler tasarım yaklaşımı sayesinde sistemin mevcut ekipmanlara minimum teknik müdahale ile entegre edilebilmesi ve bakım personelinin özel bir eğitim gerektirmeden sistemi devreye alabilmesi, pratik uygulanabilirlik açısından önemli avantajlar sunmaktadır.

Bununla birlikte, sistemin bazı kısıtlamaları da mevcuttur. Yapay sinir ağı modelinin eğitimi için kapsamlı veri seti gereksinimi, yüksek tork değerlerinin doğrulama ihtiyacı ve farklı endüstriyel haberleşme protokollerinin tam entegrasyonu gibi konular gelecek çalışmalarda ele alınması gereken önemli hususlardır. Gelecek araştırmalarda adaptif öğrenme algoritmaları ile sistemin kendini kalibre etme yeteneğinin geliştirilmesi, edge computing uygulamaları ile veri işleme kapasitesinin artırılması ve prediktif bakım fonksiyonlarının implementasyonu hedeflenmektedir.

Geliştirilen sistemin endüstriyel üretim süreçlerine potansiyel etkileri göz önüne alındığında, kalite kontrol süreçlerinin dijitalleştirilmesi ve otomatikleştirilmesi, üretim hatalarının erken tespiti ve önlenmesi, veri tabanlı karar verme süreçlerinin güçlendirilmesi ve Endüstri 4.0 dönüşümünün hızlandırılması gibi önemli katkılar sağlayacağı öngörülmektedir. Bu bulgular, geliştirilen sistemin endüstriyel modernizasyon çalışmalarında etkili bir çözüm olarak konumlandırılabilirliğini göstermektedir. Özellikle maliyet etkinliği ve kolay uygulanabilirlik özellikleri, sistemin geniş çaplı endüstriyel adaptasyonu için potansiyel sunmaktadır.

KAYNAKLAR

- [1] M.-C. Marica, N. Bizon, and I. Bostan, "Upgrading Legacy Automation Equipment to Achieve Industry 4.0 Compatibility," Jun. 2023, doi: 10.1109/ecai58194.2023.10194152.
- [2] "Industry 4.0: A Review of Digital Retrofitting Solutions for Legacy Manufacturing Systems," Nov. 2022, doi: 10.3233/atde220557.

International Open Source Conference (UAKK) 2024

- [3] Z. Huang, C. Jowers, D. Kent, A. Dehghan-Manshadi, and M. S. Dargusch, "The implementation of Industry 4.0 in manufacturing: from lean manufacturing to product design," *The International Journal of Advanced Manufacturing Technology*, Jun. 2022, doi: 10.1007/s00170-022-09511-7.
- [4] T. H. Kwok and T. Gaasenbeek, "A production interface to enable legacy factories for industry 4.0," *Engineering research express*, Sep. 2023, doi: 10.1088/2631-8695/acfec9.
- [5] M. Słowik and H. Sierocka, "Manufacturing equipment retrofitting towards Industry 4.0 standards — a systematic overview of the literature," *Engineering Management in Production and Services*, Sep. 2023, doi: 10.2478/emj-2023-0017.
- [6] D. Patel and P. Kalgutkar, "Predictive Maintenance for Industrial Equipment Using Machine Learning," *International Journal of Advanced Research in Science, Communication and Technology*, Aug. 2024, doi: 10.48175/ijarsct-19379.
- [7] A. Ghavidel, H. Park, S. Kovacic, and A. Sousa-Poza, "A New Predictive Maintenance Approach: novel integration of GMM-LSTM for prediction of latent state and failure location of Rotating Machinery," Jun. 2024, doi: 10.1109/sose62659.2024.10620926.
- [8] T. Harsh, T. Kumar, A. Mohanty, and A. Pandey, "Predictive Maintenance of Industrial Machines using ML and IoT," May 2024, doi: 10.1109/aiiot58432.2024.10574756.
- [9] J. N. A. Malaiyappan, G. Krishnamoorthy, and S. Jangoan, "Predictive Maintenance using Machine Learning in Industrial IoT," *International journal of innovative science and research technology*, Apr. 2024, doi: 10.38124/ijisrt/ijisrt24mar984.
- [10] O. B. Olorienu and J. O. Ogala, "Integrating artificial intelligence and mathematical models for predictive maintenance in industrial systems," *Fudma Journal of Sciences*, Jun. 2024, doi: 10.33003/fjs-2024-0803-2593.

UAKK 2024

Açık Kaynaklı Makine Öğrenmesi Yöntemleri Kullanarak Gömülü Sistemler ile Güneş Panellerinde Toz Tespiti

Solar Panel Dust Detection Using Embedded Systems with Open Source Machine Learning Methods

Muhammed Emin Yerli
Department of Electrical Engineering
Istanbul Technical University
İstanbul, Türkiye
yerlim22@itu.edu.tr

Aydın Tarık Zengin
Department of Electrical Engineering
Istanbul Technical University
İstanbul, Türkiye
tarik.zengin@itu.edu.tr

Özetçe— Bu çalışma, güneş panelleri üzerinde oluşan kir, toz, kar ve hasar gibi dış etkenlerin, panel verimliliğine olan olumsuz etkilerinin nesne tanıma algoritmalarıyla tespit edilmesine odaklanmaktadır. Güneş panellerinde biriken bu unsurlar, enerji üretimini önemli ölçüde düşürerek ciddi verim kayıplarına yol açmaktadır. Bu bağlamda, açık kaynaklardan elde edilen 7485 görüntü kullanılarak, dört farklı sınıfa (temiz, kirli, hasarlı, karla kaplı) ayrılan geniş ve kapsamlı bir veri seti oluşturulmuştur. Veri seti üzerinde yapılan eğitimler, derin öğrenme tabanlı SSD Mobilenet ve Faster R-CNN gibi modeller kullanılarak gerçekleştirilmiştir. Bu modeller hem masaüstü bilgisayar ortamında hem de gömülü sistem platformu olan NVIDIA Jetson Nano’da eğitilmiş ve test edilmiştir. Elde edilen sonuçlar, masaüstü bilgisayarda daha yüksek işlem hızları sunarken, gömülü sistem platformunun yerinde tespitler için uygun bir alternatif olduğunu göstermektedir. Bu çalışmanın sonuçları, güneş enerjisi santrallerindeki verimlilik kayıplarını en aza indirerek bakım maliyetlerini düşürme ve panellerin uzun vadeli performansını artırma potansiyeline sahip yenilikçi bir çözüm sunduğunu göstermektedir.

Anahtar Kelimeler — Güneş paneli, nesne tanıma, derin öğrenme, Jetson Nano, verimlilik.

Abstract— This study focuses on detecting the adverse effects of external factors such as dirt, dust, snow, and damage on solar panel efficiency using object recognition algorithms. These external elements that accumulate on solar panels significantly hinder energy production, leading to major efficiency losses. A comprehensive dataset of 7485 images, collected from open sources, was created and categorized into four distinct classes (clean, dirty, damaged, snow-covered). Training of deep learning-based models, such as SSD Mobilenet and Faster R-CNN, was conducted on both a desktop computer and the NVIDIA Jetson Nano embedded system platform. The results indicate that while the desktop computer provides higher processing speeds, the embedded system platform is a suitable alternative for on-site detection. The findings suggest that this innovative solution has the potential to minimize efficiency losses in solar power plants, reduce maintenance costs, and enhance the long-term performance of solar panels.

Keywords — Solar panel, object detection, deep learning, Jetson Nano, efficiency.

I. GİRİŞ

Güneş enerjisi, son yıllarda dünyada en hızlı büyüyen yenilenebilir enerji kaynaklarından biri haline gelmektedir. 2021 yılında, küresel yenilenebilir üretim kapasitesi 257 GW artarak toplamda 3064 GW’a ulaşmıştır [1]. Güneş enerjisi, 133 GW (%19) artışla kapasite genişlemesine öncülük etmiş ve bu sayede küresel güneş kapasitesi 2021’de 843 GW’a yükselmiştir [1][2]. Örneğin, Çin, 2020 yılında yıllık 253 GW kapasitesinden 307 GW ekleyerek en büyük katkıyı yapmıştır. Öte yandan, ABD, güneş enerjisi kapasitesini %27 oranında artırarak 94 GW seviyesinde güneş enerjisi üretimine ulaşmıştır [1].

Güneş enerjisi, karbon salınımını sıfıra indirerek iklim değişikliğiyle mücadelede önemli bir rol oynamakta ve fosil yakıt bağımlılığını azaltmaktadır. 2045 yılına kadar tamamen yenilenebilir enerjiye geçiş hedefleri, birçok ülkenin stratejik enerji planları arasında yer almaktadır [3]. Güneş enerjisi yalnızca çevre dostu olmakla kalmamakta, aynı zamanda geniş arazilerde kurulan büyük ölçekli santrallerde maliyet avantajı da sunmaktadır. Ancak, güneş panellerinin verimliliği, çevresel etkenlerden doğrudan etkilenmektedir. Panellerde biriken toz, kir ve kar gibi unsurlar, güneş ışığının panellere ulaşmasını engelleyerek enerji üretim kapasitesini düşürmektedir [4][5].

Özellikle çöl ve endüstriyel bölgelerde kurulu paneller, toz birikimi nedeniyle büyük enerji kayıplarına uğrar. Orta Doğu ve Kuzey Afrika gibi yüksek tozlanma riskine sahip bölgelerde kurulu panellerde, verimlilik %35 oranında azalabilmektedir [8]. Hindistan’da ise bir yıl boyunca temizlenmeyen panellerde %50’ye varan enerji kaybı gözlemlenmiştir [4][5]. En kötü senaryolarda, bazı sanayi bölgelerinde %80’e kadar performans kaybı rapor edilmiştir [5]. Bu kayıplar, büyük ölçekli güneş enerjisi santrallerinde ciddi ekonomik zararlara yol açmakta ve yenilenebilir enerji üretim hedeflerini tehdit etmektedir. Ayrıca, temizleme ve bakım işlemlerinin düzenli ve maliyet etkin bir şekilde yapılmaması hem işletme maliyetlerini artırmakta hem de panellerin ömrünü olumsuz yönde etkilemektedir.

Bu çalışmanın temel amacı, güneş panellerinde oluşan kirlilik ve fiziksel hasarların otomatik tespiti için bir yazılım modeli geliştirmektir. Amaç, çevresel faktörlerin güneş panelleri üzerindeki olumsuz etkilerini en aza indirirken, enerji üretim kayıplarını tespit eden ve bakım süreçlerini optimize eden bir çözüm sunmaktır. Bu doğrultuda, geliştirilecek yazılımın hem Windows işletim sistemine sahip masaüstü bir bilgisayarda hem de düşük enerji tüketimi gerektiren NVIDIA Jetson Nano gibi gömülü platformlarda sorunsuz çalışması hedeflenmektedir.

Bu çalışma kapsamında, güneş panelleri ve makine öğrenimi alanlarında yapılan literatür taramasıyla, son 20 yılda toz ve kir sorununa odaklanan araştırmalar analiz edilmiştir. Çalışmalar görüntü işleme ve gömülü sistemler başlıkları altında iki ana filtre ile sınıflandırılmıştır. Toplamda 50 adet makale belirlenmiş ve bu makalelerin yıllara göre dağılımı analiz edilmiştir. Özellikle 2017 sonrasında bu alanda yapılan çalışmaların sayısında önemli bir artış olduğu gözlemlenmiştir. Bu durum, güneş panellerindeki verim kaybının dünya genelinde giderek büyük bir sorun haline gelmesiyle ilişkilendirilebilir.

Görüntü işleme teknikleri, güneş panellerinde toz tespiti çalışmalarında yaygın olarak kullanılmaktadır. 50 makale arasından seçilen 22'sinde, derin öğrenme ve makine öğrenmesi modelleri öne çıkmaktadır. Bu çalışmalarda CNN (Convolutional Neural Networks), ResNet (Residual Networks) ve Yolo (You Only Look Once) gibi modellerin yanı sıra, çeşitli CNN mimarileri (örneğin LeNet, ResNet, MobileNet) kullanılmıştır. Bu modellerin çevresel faktörlerin görüntü üzerinden tespitinde yüksek performans sergilediği gözlemlenmiştir. Ayrıca, ANN (Artificial Neural Network), RNN (Residual Neural Network) ve LSTM (Long Short-Term Memory) gibi diğer yöntemler de farklı problem türlerinde etkili çözüm önerileri sunmuştur. Bu algoritmaların çoğu Keras ve TensorFlow kütüphaneleri ile desteklenmiştir.

Literatür taramasında güneş panellerinde toz tespiti için kullanılan görüntü işleme metotları incelenirken, gömülü sistemler üzerindeki uygulamalar ayrıca ele alınmıştır. Gömülü sistemler, düşük enerji tüketimi ve gerçek zamanlı izleme gibi avantajlar sunduğu için önemli bir çözüm olarak öne çıkmaktadır. Ancak, incelenen 50 makaleden yalnızca bir tanesi, gömülü platformda görüntü işleme uygulamıştır. Bu çalışmada, NVIDIA Jetson platformu ile güneş paneli yüzeylerindeki toz birikimlerinin tespiti ele alınmıştır [23].

II. VERİ SETİNİN HAZIRLANMASI VE MODEL EĞİTİMİ

Bu bölümde güneş panellerinde toz tespiti için kullanılan görüntü işleme modellerinin eğitim süreci sırasıyla açıklanacaktır.

- Eğitim ve nesne tanıma işlemleri için masaüstü bilgisayara gerekli kurulumların yapılması
- Eğitimde kullanılacak veri setinin hazırlanması ve düzenlenmesi, bu veri setinin eğitim ve test verisi olarak ayrılması
- Eğitim için en uygun modellerin seçilmesi
- Masaüstü bilgisayarda eğitimin tamamlanması

A. Masaüstü Bilgisayar Kurulumlar

Model eğitimi, Tablo 1'de özellikleri belirtilen masaüstü bilgisayar üzerinde gerçekleştirilmiştir:

TABLE I. KULLANILAN BİLGİSAYARIN ÖZELLİKLERİ

Özellik	Detay
İşlemci	Intel Core i5-14600K
Ekran Kartı	NVIDIA RTX 4070 SUPER 12GB
Bellek	32 GB DDR5 7000 MHz
İşletim Sistemi	Windows 11 Pro

TensorFlow eğitimi için gerekli kurulumlar, aşağıda listelenen yazılımlar ile sağlanmıştır:

- Anaconda Navigator 1.9.12
- PyCharm Community Edition 2024.1.1
- Cuda 10.0.130
- cuDNN 10.0 v7.4.2.24
- LabelImg v1.8.1

B. Veri Setinin Hazırlanması

Makine öğrenimi ve derin öğrenme modellerinin doğru bir şekilde eğitilebilmesi için, yüksek kaliteli ve doğru şekilde etiketlenmiş bir veri setine ihtiyaç vardır. Bu projede, tüm kaynaklar taranmış ve eğitim için en uygun veriler, geniş bir veri tabanına ve çeşitliliğe sahip olan açık kaynak veri sağlayıcısı Kaggle üzerinden temin edilmiştir. Kaggle, araştırmacılar ve veri bilimciler tarafından sıklıkla tercih edilen bir platform olduğundan, burada projenin ihtiyaçlarına uygun, geniş kapsamlı görüntülere ulaşmak mümkün olmuştur.

Kaggle içerinden 8 farklı kaynaktan 12567 adet görüntü temin edilmiştir. Bu görüntüler, güneş panellerinin çeşitli koşullardaki hallerini yansıtmaktadır. Kaynaklardan gelen verilerin bazıları sınıflandırılmamış olup, bazıları ise kendi içinde gibi kategorilere ayrılmıştır.

Tüm verileri bu proje kapsamında Şekil 1'de gösterilen dört ana sınıf altında sınıflandırılmıştır. Bu sınıflar: Clean (Temiz), Dirty (Kırlı), Damaged (Hasarlı) ve Snow-Covered (Karla Kaplı)



Fig. 1. Clean, Dirty, Damaged ve Snow-Covered sınıflarına ait güneş paneli örnekleri.

Veri setini düzenlerken aynı açıdan birden fazla kez çekilmiş ve gereksiz yere tekrarlayan görüntüler silinmiş, veri setinin homojenliği ve çeşitliliği artırılmıştır. Bu düzenlemeler sonucunda her sınıfa ait görüntü sayısı Tablo 2'de, her sınıfa ait toplam nesne sayısı Tablo 3'de gösterilmektedir. Bazı görüntülerde birden fazla nesne olduğu için toplam görüntü sayısı ile toplam nesne sayısı birbirinden farklı çıkmaktadır.

TABLE II. HER SINIFA AIT TOPLAM GÖRÜNTÜ SAYISI

Clean	Dirty	Damaged	Snow Covered	Toplam Görüntü
4148	2940	274	123	7485

TABLE III. HER SINIFA AIT TOPLAM NESNE SAYISI

Clean	Dirty	Damaged	Snow Covered	Toplam Nesne
5271	3241	524	166	9202

Tüm veriler sınıflandırıldıktan sonra, masaüstü bilgisayara yüklenen Labelmg programı ile her bir görüntü içerisindeki nesne, kendi sınıfına göre gibi etiketlenmiştir. Bazı görüntülerde birden fazla nesne veya farklı sınıflarda nesnelere bulunmaktadır. Bu duruma örnek Şekil 2'deki gibi hem 'Clean' hem de 'Dirty' sınıfında nesneye sahip görüntü doğru bir şekilde etiketlenmiştir.

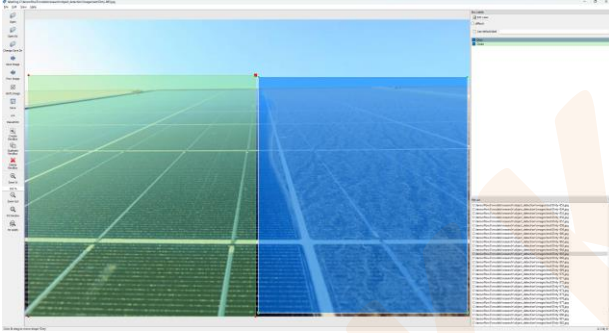


Fig. 2. Hem 'Clean' hemde 'Dirty' sınıfında bir nesneye sahip görüntünün labelmg programı ile etiketlenmesi

Eğitim öncesi son adım olarak veri seti, 2 farklı klasöre %80'i eğitim verisi %20'si test verisi olacak şekilde ayrılması gerekmektedir. Başarılı bir eğitim için Tablo 4'de görüldüğü üzere, toplam verinin %80,2'si eğitim verisi olacak şekilde, %19,8'i ise test verisi olacak şekilde ayrılmıştır.

TABLE IV. VERİ SETİNİN TEST VE TRAIN KLASÖRÜNE DAĞILIMI

	Clean	Dirty	Damaged	Snow Covered	Toplam
Toplam	4148	2940	274	123	7485
Train	3348	2340	214	103	6005
Test	800	600	60	20	1480
Oran (%) Train/Toplam	80,7	79,5	78,1	83,7	80,2

C. Model Eğitimi

Bu projede kullanılan derin öğrenme modelleri, TensorFlow'un "TensorFlow Detection Model Zoo" içerisinde yer alan önceden eğitilmiş modeller arasından seçilmiştir. Model seçimi yapılırken, özellikle COCO (Common Objects in Context) veri seti üzerinde eğitilmiş olan ve farklı hız (ms) ve doğruluk oranlarına sahip olan COCO eğitilmiş model grubu tercih edilmiştir. Model seçiminde hız ve doğruluk performanslarındaki farklılıklar en önemli kriterlerdir. Bu çeşitlilik, farklı görevlerde ve donanımlarda performans analizlerinin kapsamlı yapılmasını sağlamıştır.

1) Faster R-CNN Inception V2 ile Eğitim

Seçilen bu modelin Speed(ms) değeri 58 ve COCO mAP değeri ise 28'dir ve orta hız ve doğruluk sunan bu model, genel denge sağlamak amacıyla tercih edilmiştir. Model ile eğitim masaüstü bilgisayarda başarıyla tamamlanmıştır. Eğitim 200 bin adım sürmüştür ve her bir adım ortama 1 saniyeden, toplamada eğitim 2 gün 7 saat 30 dakika sürmüştür.

2) SSD Mobilenet V1 ile Eğitim

Seçilen bu modelin Speed(ms) değeri 30 ve COCO mAP değeri ise 21'dir ve yüksek hız ve düşük doğruluk sunan bu model, gerçek zamanlı uygulamalarda kullanılmak üzere seçilmiştir. Gömülü sistemlerde kullanılmaya daha uygundur. Model ile eğitim masaüstü bilgisayarda başarıyla tamamlanmıştır. Eğitim 200 bin adım sürmüştür ve her bir adım ortama 5 saniyeden, toplamada eğitim 11 gün 14 saat 30 dakika sürmüştür.

3) ResNet Mask R-CNN Inception ResNet V2 ile Eğitim

Seçilen bu modelin Speed(ms) değeri 771 ve COCO mAP değeri ise 36'dir ve doğruluğun önemli olduğu ve detaylı analiz gerektiren uygulamalar için seçilmiştir. Fakat eğitim başarı ile tamamlanamamıştır.

Bunun sebebi ilk iki modelin "Boxes" olarak çıktı vermesi fakat bu modelin "Masks" olarak çıktı vermesi olarak tespit edilmiştir. Bu yüzden "Boxes" olarak çıktı veren yeni bir model seçilmiştir.

a) RFCN ResNet101 ile Eğitim

ResNet V2 Mask R-CNN Inception yerine seçilen bu modelin Speed(ms) değeri 92 ve COCO mAP değeri ise 30'dir. Bu model ile eğitim masaüstü bilgisayarda başarıyla tamamlanmıştır. Eğitim 200 bin adım sürmüştür ve her bir adım ortama 2 saniyeden, toplamada eğitim 4 gün 15 saat 33 dakika sürmüştür.

4) Faster R-CNN NAS ile Eğitim

Seçilen bu modelin Speed(ms) değeri 1833 ve COCO mAP değeri ise 43'dir ve en yüksek doğruluk oranına sahip olan bu model, doğruluğun ön planda olduğu durumlarda tercih edilmiştir. Model ile eğitim masaüstü bilgisayarda başarıyla tamamlanmıştır. Eğitim 50 bin adım sürmüştür ve her bir adım ortama 17 saniyeden, toplamada eğitim 9 gün 20 saat 36 dakika sürmüştür.

III. DENEYSEL SONUÇLAR

Dört farklı model ile eğitilen sistemlerin, masaüstü bilgisayarda ve Jetson Nano platformunda nasıl performans gösterdiğini analiz etmek amacıyla, eğitim verisi haricinde 80 adet farklı boyutlarda ve farklı ortam koşullarında çekilmiş test görüntüsü ile nesne tanıma uygulamaları yapılmıştır.

A. Masaüstü Bilgisayar Üzerinde Nesne Tanıma Sonuçları

Yapılan testler sonucunda masaüstü bilgisayar için nesne tanıma başarı oranları Tablo 5'de, tek bir görüntü için ortalama nesne tanıma süreleri ise Tablo 6'de gösterilmektedir. Bu sonuçlara göre en başarılı model RFCN ResNet101 modelidir, en hızlı çalışan model ise SSD Mobilenet V1 modelidir.

TABLE V. MASAÜSTÜ BİLGİSAYARDA NESNE TANIMA BAŞARI ORANLARI

Modeller	Clean %	Dirty %	Damaged %	Snow Covered %	Toplam Başarı Oran %
Faster R-CNN Inception V2	79.1	84.8	18.2	55.9	59.47
SSD Mobilenet V1	82.5	8.4	18.5	9.4	29.69
RFCN ResNet101	83.5	86.7	32.9	41.3	61.08
Faster R-CNN NAS	96.3	28.4	0	0	31.17

TABLE VI. MASAÜSTÜ BİLGİSAYARDA TEK BİR GÖRÜNTÜ İÇİN ORTAMA NESNE TANIMA SÜRELERİ

Modeller	Nesne Tanıma Süresi
Faster R-CNN Inception V2	401 ms
SSD Mobilenet V1	43 ms
RFCN ResNet101	535 ms
Faster R-CNN NAS	1320 ms

En başarılı model olan RFCN ResNet101 COCO ile nesne tanıma örnekleri her bir sınıf için Şekil 3'de gösterilmektedir.

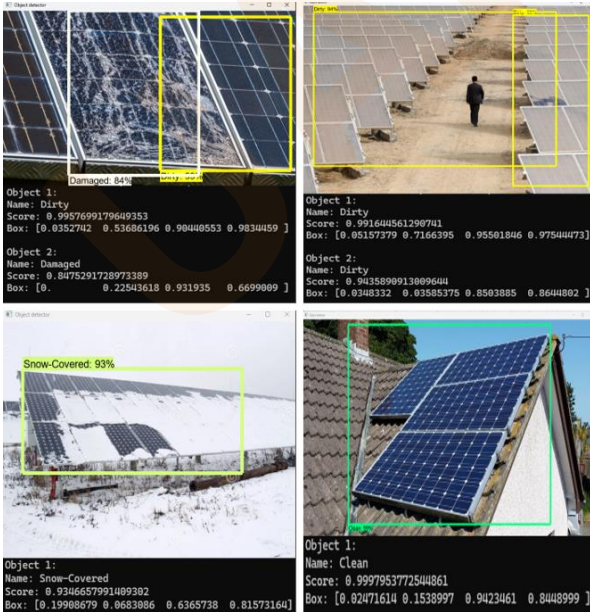


Fig. 3. Masaüstü bilgisayar üzerinden nesne tanıma örnekleri.

B. Jetson Nano Üzerinde Nesne Tanıma Sonuçları

Jetson Nano üzerinde başlatılan nesne tanıma işlemlerinde, dört modelden yalnızca biri başarıyla çalıştırılabilirdi. Kalan üç modelde nesne tanıma başarısız olmuştur. Bu modeller:

- Faster R-CNN Inception V2
- RFCN ResNet101
- Faster R-CNN NAS

Bu üç modelin test edilmesi sırasında, Jetson Nano'da görülen "System throttled due to Over-current" ve "Low Memory Warning" hataları ile karşılaşıldı. Modellerin yüksek bellek ve işlem gücü gereksinimleri nedeniyle Jetson Nano, işlemleri tamamlayamadan durdu. Bu üç modelde nesne tanıma işlemi başarıyla gerçekleştirilemedi.

Diğer modellerin aksine SSD Mobilenet V1 modeli ile tüm test görüntüleri üzerinde nesne tanıma işlemi başarıyla tamamlandı. Bu modelin seçilme sebeplerinden biri, hafif ve hızlı olmasıdır. Ancak modelin bu avantajı, doğruluk oranının düşük olmasına neden oldu. Nesne tanıma işlemi sonucunda, görüntülerdeki güneş panelleri doğru şekilde tespit edilse de tüm nesnelere yanlış sınıflandırıldı.

Bu durumlara örnek olarak Şekil 4'de, farklı çözünürlükteki 2 görüntü için nesne tanıma çıktıları verilmiştir.



Fig. 4. Jetson Nano üzerinden hatalı nesne tanıma örnekleri.

IV. SONUÇ ÖNERİLER

A. Sonuç

Bu çalışmada, güneş panellerinde toz tespiti amacıyla çeşitli derin öğrenme modelleri kullanılarak nesne tanıma işlemleri gerçekleştirilmiştir. Model performansları, hem masaüstü bilgisayarda hem de gömülü bir sistem olan NVIDIA Jetson Nano üzerinde değerlendirilmiş, bu iki platformun tanıma süreleri ve başarı oranları karşılaştırılmıştır. Çalışmada Faster R-CNN Inception V2, SSD Mobilenet V1, RFCN ResNet101 ve Faster R-CNN NAS gibi farklı derin öğrenme modelleri kullanılmış ve sonuçlar analiz edilmiştir.

Masaüstü bilgisayarda elde edilen sonuçlar, yüksek performanslı bir donanımda çalıştırılan modellerin, özellikle daha karmaşık ve derin yapılar içeren modellerde daha iyi başarı oranları sunduğunu göstermiştir. Ancak Jetson Nano gibi sınırlı kaynaklara sahip gömülü sistemlerde, modellerin performansının ciddi şekilde düştüğü gözlemlenmiştir. Örneğin, SSD Mobilenet modelinde Jetson Nano üzerinde tanıma süresi 360 saniye gibi uzun bir süreye çıkmış, diğer modellerde ise tamamen başarısızlıkla sonuçlanmıştır. Bu da gömülü sistemlerde kullanılacak modellerin hafif ve optimize edilmiş yapılar olmasının gerekliliğini ortaya koymaktadır.

B. Öneriler

Model Optimizasyonu: Jetson Nano gibi sınırlı donanım kaynaklarına sahip cihazlarda daha verimli sonuçlar elde edebilmek için, kullanılan derin öğrenme modellerinin donanıma uygun şekilde optimize edilmesi gerekmektedir. Özellikle TensorRT gibi araçlarla model optimizasyonu yapılarak tanıma süreleri kısaltılabilir ve başarı oranları artırılabilir.

Donanım Yükseltmesi: Daha karmaşık derin öğrenme modelleri kullanmak isteyenler için, Jetson Nano gibi düşük güç tüketimine sahip cihazlar yerine, Jetson Xavier gibi daha güçlü gömülü sistemler tercih edilebilir. Bu sayede hem daha hızlı hem de daha başarılı sonuçlar elde edilebilir.

Veri Seti Genişletilmesi: Çalışmada kullanılan veri setinin genişletilmesi ve çeşitlendirilmesi, özellikle tozun farklı derecelerini ve diğer kirlilik türlerini kapsayacak şekilde geliştirilmesi, modelin başarımını artırabilir. Ayrıca, gerçek zamanlı görüntü işleme yapılacaksa, bu veri setlerinin gerçek zamanlı etiketleme yöntemleriyle güncellenmesi faydalı olacaktır.

Hafif Modellerin Kullanımı: Özellikle mobil ve gömülü sistemler için geliştirilen EfficientDet, MobileNetV2 gibi daha hafif ve hızlı modellerin kullanılması, gömülü sistemlerdeki performansı artıracaktır. Gelecekte yapılacak çalışmalar bu modellerin kullanımını daha detaylı inceleyebilir.

KAYNAKLAR

- [1] Md Saif Hassan Onim, Zubayar Mahatab Md Sakif ve Ali M.S., "SolNet: A Convolutional Neural Network for Detecting Dust on Solar Panels", 2022.
- [2] Tonatihu Cruz-Rojas, Jesus Alejandro Franco, "A novel comparison of image semantic segmentation techniques for detecting dust in photovoltaic panels using machine learning and deep learning", 2023.
- [3] Krishna Sandeep Ayyagari, Yuvraj Munian, "Simultaneous Detection and Classification of Dust and Soil on Solar PhotoVoltaic Arrays Connected to A Large-Scale Industry: A Case Study", 2024.
- [4] Dania Saquib, M.N.N., "Image Processing Based Dust Detection and prediction of Power using ANN in PV systems", 2022.
- [5] Hassan Qasem, A.M. ve Banda, P., "Assessing Dust on PV Modules Using Image Processing Techniques", 2016.
- [6] Atl Emre COŞGUN, H.D., "The experimental study of dust effect on solar panel efficiency", 2022.
- [7] Kamal Adel Abuqaad, A.F., "A Novel Technique for Detecting and Monitoring Dust and Soil on Solar Photovoltaic Panel", 2020.
- [8] Mostafa. F. Shaaban, Amal Alarif, M.M.U.T.A.H.O. ve Al-Ali, A.R., "A New Data-Based Dust Estimation Unit for PV Panels", 2020.
- [9] Sachin Mehta, Amar P. Azad, S.A.C.V.R. ve Kalyanaraman, S., "DeepSolarEye: Power Loss Prediction and Weakly Supervised Soiling Localization via Fully Convolutional Networks for Solar Panels", 2018.
- [10] Ritu Maity, M.S.A. ve Pati, A., "An Approach for Detection of Dust on Solar Panels Using CNN from RGB Dust Image to Predict Power Loss", 2020.
- [11] Siyuan Fan, Yu Wang, S.C.B.Z.T.S.P.L., "A deep residual neural network identification method for uneven dust accumulation on photovoltaic (PV) panels", 2021.
- [12] Shengxian Cao, Xinyue Ma, S.F.T.W., "IDS-Net: Integrated Network for Identifying Dust State of Photovoltaic Panels", 2021.
- [13] Weerasekara, W. ve Maduranga, M., "A Deep Learning-Based Approach for Detecting Dust on Solar Panels", 2023.
- [14] S. Prabhakaran1, R.A.U. ve Preetharoselyn, J., "Deep Learning-Based Model for Defect Detection and Localization on Photovoltaic Panels", 2022.
- [15] Yichuan Shao, Can Zhang, L.X.H.S.Q.Z.L.Z., "A new dust detection method for photovoltaic panel surface based on Pytorch and its economic benefit analysis", 2024.
- [16] Harri Barrera, Kimberley Weiss, G.P.W.R., "Transfer Learning in Dust Detection on Solar Panels", 2023.
- [17] Dhruv Ferrell, E.A., "Adapting and Generalizing Convolutional Neural Networks in Detecting Dust on Solar Panels", 2023.
- [18] Uditha, B.D.I., "Identify dusts on solar panels using drone still images", 2019.
- [19] Marek Jaszczur, Ambalika Koshti, W.N. ve Şedor, P., "An investigation of the dust accumulation on photovoltaic panels", 2014.
- [20] Yuan Gao, S.L., "A deep learning-based method detects dust from solar PV panels through Unmanned Aerial Vehicles, Journal of Physics", 2023.
- [21] Massaab El Ydrissi, Hicham Ghennioui, E.G.B.A.A.M.A.I.T.A.F., "Dust InSMS: Intelligent soiling measurement system for dust detection on solar mirrors using computer vision methods", 2023.
- [22] Fuhao Sun, Cheng Yang, H.C.Z.L.J.S.B.Z.K.H., "Dust Detection Techniques for Photovoltaic Panels from a Machine Vision Perspective: A Review", 2023.
- [23] Nazmun Nahar Karima, Kamruzzaman Rimon, M.S.M.M.H. ve Bhuyan, M.H., "Advanced Image Processing Based Solar Panel Dust Detection System", 2023.

Diagnosis Enhancement Using Deep Learning: YOLOv11-YOLOv8 for Alzheimer's and Parkinson's Disease Detection

Bilge Koç
Department of Computer Engineering
Malatya Turgut Özal University
Malatya, Türkiye
kocbilgeec@gmail.com

Dilber Çetintaş
Department of Computer Engineering
Malatya Turgut Özal University
Malatya, Türkiye
dilber.cetintas@ozal.edu.tr

Abstract— Object detection is a crucial image processing technique, particularly in the field of healthcare. It plays a vital role in the accurate identification of diseases and the effective implementation of treatments. Given the brain's complex anatomical structure and the nature of neurological disorders, there is an increasing need for accurate diagnostic methods. Current analysis techniques may impact diagnostic accuracy due to factors such as human errors and low processing speeds, prompting healthcare professionals to seek more reliable automated diagnostic systems..

This study aims to accurately and rapidly detect neurological disorders such as Parkinson's and Alzheimer's using Magnetic Resonance (MR) images without requiring preprocessing. For this purpose, the "enfermedades_cerebro-uagpl" dataset was utilized, which was sourced from the Roboflow platform and includes MR images of Alzheimer's, Parkinson's, and normal brains. The analyses were conducted using the latest versions of the YOLO algorithms, specifically YOLOv8 and the experimentally developed YOLOv11 models. While YOLOv11 achieved an accuracy rate of 100%, YOLOv8 demonstrated a success rate of 89.7%.

Keywords—Alzheimer, Parkinson Yolo11, Yolo8, Nörodejeneratif, Manyetik Rezonanz

I. INTRODUCTION

Dementia is a group of disorders that affect brain functions, leading to declines in memory and physical abilities. Various types of dementia manifest with different symptoms and mechanisms in specific regions of the brain. Among these, Alzheimer's and Parkinson's diseases are classified as neurodegenerative disorders, causing progressive damage to brain tissue over time [1]. Both diseases are believed to result in significant structural and functional impairments in the brain during years of silent progression before noticeable symptoms emerge [2]. According to data from the World Health Organization (WHO) [3], approximately 50 million people worldwide were diagnosed with dementia in 2019, with nearly 10 million new cases being added annually.

Parkinson's disease is a serious neurological disorder that affects the central nervous system and progresses over time. This condition manifests with symptoms such as tremors, muscle rigidity, and slowed movements due to impairments in

motor control. As the disease worsens, it significantly hampers patients' ability to perform daily activities. The primary underlying cause of Parkinson's disease is the loss of dopamine-producing cells in the brain. This cell loss leads to a reduction in dopamine levels, which are essential for coordinating movements and maintaining balance. In addition to motor dysfunction, Parkinson's disease also affects cognitive functions, potentially causing memory loss, depression, and anxiety-related psychological issues [4].

Among the various types of dementia, Alzheimer's disease is the most common and is a severe neurological condition that primarily affects elderly individuals. It leads to memory loss and disruptions in daily activities. With the global aging population steadily increasing, a significant rise in the number of Alzheimer's patients has been observed, and this trend is expected to accelerate in the coming years [5].

Magnetic Resonance (MR) imaging plays a crucial role in the diagnosis of Parkinson's and Alzheimer's diseases due to its ability to detect structural changes, such as the loss of brain cells. High-resolution MR techniques enable the non-invasive examination of neurological changes, providing healthcare professionals with valuable guidance during diagnostic and follow-up processes [6]. Advances in the application of artificial intelligence (AI) to MR imaging have reshaped diagnostic and treatment processes, offering groundbreaking solutions for the early and accurate detection of neurological disorders. As the prevalence of diseases such as Alzheimer's and Parkinson's increases with the aging population, AI's ability to identify subtle abnormalities in brain images has become a significant asset. Unlike traditional diagnostic and treatment methods, these AI-driven approaches aim to address limitations such as subjectivity, overlooked early-stage symptoms, and findings undetectable by human observers [7].

Numerous artificial intelligence models have been developed in the literature for the diagnosis of Alzheimer's and Parkinson's diseases. Shreya Reddy et al. successfully classified MR images of 221 individuals with Parkinson's disease and 221 healthy individuals using a Convolutional Neural Network (CNN) [8]. Similarly, Desai and colleagues achieved an 88% success rate by processing three-dimensional brain images with a CNN model they developed for the rapid and accurate diagnosis of Parkinson's disease [9]. Another study aimed at early diagnosis analyzed 7,635 MR images labeled as mild, normal, or Alzheimer's. The study compared machine learning methods with CNN results and concluded that CNN models are a superior option compared to traditional machine learning approaches [10].

In the context of dementia diagnosis, the role of preprocessing techniques in improving results from CNN models applied to MR images has drawn significant attention to the field of object detection. The success of the You Only Look Once (YOLO) architecture in real-time applications was a primary factor in steering this study towards this area. Table 1 presents dementia-related studies conducted using YOLO.

TABLE I. DEMANS-YOLO LITERATURE

Article	Literature		
	Data	Architecture	Results
Sara Saad Abd-Aljabar(2023)[11],	Alzheimer 300(MCI/AD/NC)	Yolo2	%98
Islam et al. (2023)[12]	Alzheimer 500 MRI images	From YOLOv3 to YOLOv7	Max YOLOv3 and YOLOv7 = %99
Amal Alshaharni (2024)[13]	Alzheimer 500 MRI images	Yolov5, Yolov8, Yolov9	Max YOLOv9 = %86
Karthikeyan (2023)[14]	Parkinson's Drawings 210 images	Yolov8-x	%94
Hui Li(2024) [15]	582 MRI images	Yolov5 + attention	%98

II. METHOD

YOLO, short for "You Only Look Once," is a high-performance deep learning algorithm for object detection [16]. Its primary function is to predict the bounding box coordinates of objects in an image and the probabilities of their corresponding classes [17]. By analyzing the entire image in a single processing stage, YOLO achieves exceptional speed, making it particularly advantageous for applications in the healthcare domain where quick and accurate results are essential.

A. Dataset

The "enfermedades_cerebro-uagpl" dataset, published on the Roboflow platform, was utilized in this study. As shown in Figure 1, the dataset contains images categorized into three classes: Alzheimer, Parkinson, and Normal. The images are provided at a resolution of 640x640 pixels..

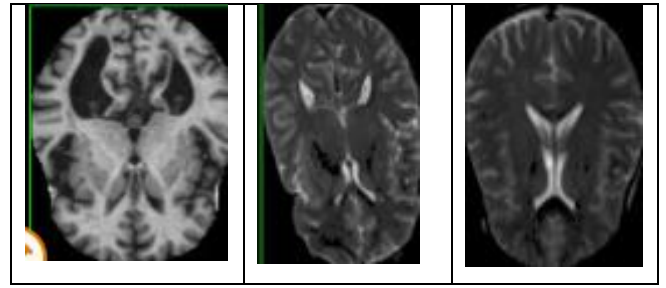
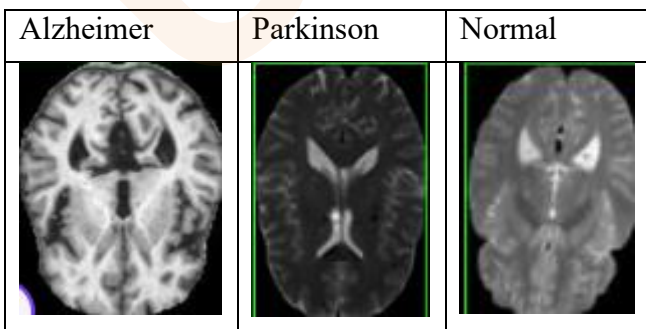


Fig. 1. Example of Dataset

The dataset consists of a total of 999 MR images, with 799 images allocated for training, 100 for validation, and 100 for testing purposes.

B. YOLO Architecture

The YOLO (You Only Look Once) algorithm performs object recognition in a single step through a neural network. During this process, the image is divided into a grid structure. If the center of an object is located within a grid cell, the algorithm determines the object's width, height, and class label for that cell. This method allows YOLO to achieve fast and efficient object detection by processing the image only once, in contrast to sliding window techniques [18]. Since its introduction in 2016, several versions have been developed to address evolving challenges and provide solutions. This study focuses on YOLO11, the latest version, and investigates its effectiveness in detecting dementia conditions.

C. YOLOv11

YOLOv11 represents the latest version of deep learning models developed for object detection. It performs object detection in a single step, offering significant improvements in both speed and accuracy compared to previous versions [20]. YOLOv11 replaces the C2f block from YOLOv8 with the more efficient C3k2 structure introduced in YOLOv7 [21]. By analyzing the entire image in a single processing cycle, YOLOv11 minimizes processing time, making it particularly suitable for real-time medical applications [22]. Additionally, YOLOv11 demonstrates superior performance in detecting small objects and achieving high classification accuracy.[23]

D. YOLOv8

In 2023, the Ultrasonic team developed an approach that combined the advantages of various object detectors to make object detection processes faster and more accurate [24]. The backbone network used for feature extraction retains a structure similar to previous versions. However, cross-connections were introduced at different stages of the network, significantly improving information flow and feature reuse. YOLOv8 introduced semantic segmentation capabilities, enabling the model to classify every pixel in an image. Additionally, the model was designed to support various tasks such as pose estimation, image classification, and directed object detection [21]. This approach resulted in richer representations by merging highly diverse feature maps. Furthermore, the reduction in the number of

hyperparameters in YOLOv8 simplified the training process considerably.

E. Evaluation Metrics

Determining the conditions under which a model achieves the best results is possible through specific metrics. Some of the metrics used include Precision, Recall, F1-Score, Accuracy, Average Precision (AP), and Mean Average Precision (mAP). Based on the confusion matrix, values such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are derived. These values are then formulated to evaluate the models. In this study, Accuracy, AP, and mAP metrics were utilized during the evaluation phase.

Accuracy:

It is the ratio of correct predictions to the total number of samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Average Precision: It is a method used to evaluate the performance of multi-class problems in object detection. The calculation considers the total area under the precision and recall curves for each object class [19].

$$AP = \int_0^1 P(R)dR \quad (2)$$

Fig. 2. **Mean Average Precision:** The mAP value is calculated by averaging the AP values computed individually for each object class. It serves as a useful metric for a more general evaluation.

$$mAP = \frac{1}{N} * \sum_{i=1}^N AP_i \quad (3)$$

III. RESULTS

This study aims to detect neurological diseases from brain Magnetic Resonance (MR) images using YOLO11, one of the latest versions of object detection algorithms. It compares the performance of YOLO8 and YOLO11 models to evaluate their success in identifying neurological diseases in brain MR images.

TABLE II. TRAINING PARAMETERS

Model	Training Setting		
	Learning Rate	Batch Size	Epoch
Yolov8	0.0005	16	50
Yolov11	0.001	16	50

The training configurations of the models used in the study are presented in detail in Table 2. The batch size and the number of epochs were kept identical during the training process, allowing for a more objective comparison of the models' learning capacities. However, the learning rates were selected based on the conditions that yielded the best results for each

model. This indicates that the parameters necessary for both models to perform optimally were meticulously optimized.

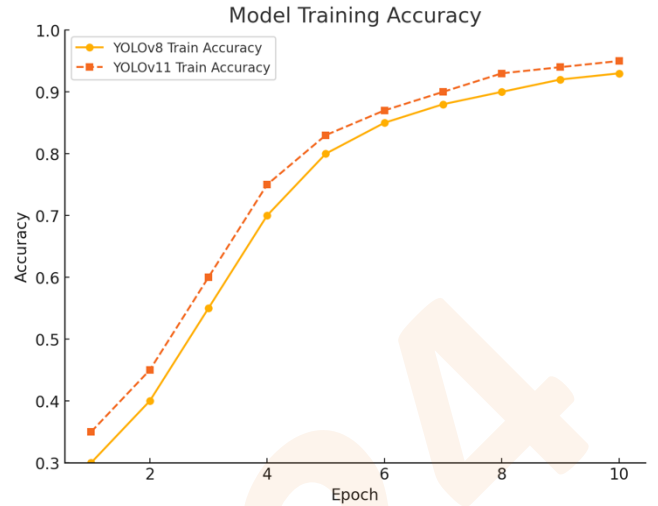


Fig. 2. Training Accuracy for YOLOv8 and YOLOv11 Models

The YOLOv11 model demonstrates a higher initial accuracy (35%) compared to the YOLOv8 model (30%) at the beginning of training. This indicates that the YOLOv11 model has a faster learning capability in the early stages of data training.

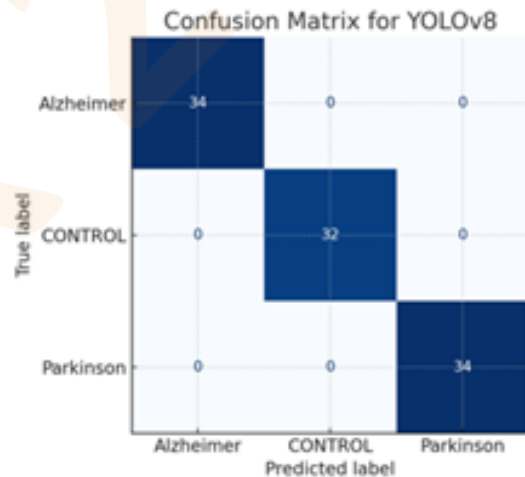


Fig. 3. Complexity Matrix for YOLOv8

Figure 3 displays the confusion matrix for the YOLOv8 model, while Figure 4 illustrates it for the YOLOv11 model. These graphs visualize how accurately each model classifies the Alzheimer, CONTROL, and Parkinson classes. While YOLOv8 results in lower error rates across all classes, the YOLOv11 model demonstrates greater generalization capability. Despite its low-level errors, the overall performance of YOLOv11 is more valuable in terms of data balance across the classes.

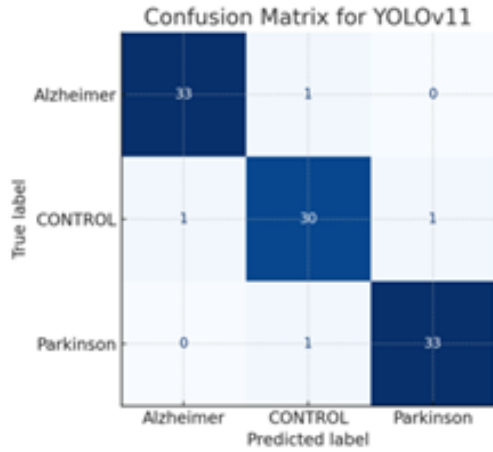


Fig. 4. Complexity Matrix for YOLOv8

While YOLOv11 achieves high accuracy in the Alzheimer and Parkinson classes, it exhibits more noticeable errors in the CONTROL class. Specifically, YOLOv11 misclassified one Parkinson patient as part of the CONTROL group and incorrectly identified two CONTROL group patients, one as Alzheimer and the other as Parkinson.

The impact of the model parameters was analyzed, and the values obtained are presented in Table 3.

TABLE III. RESULTS OF YOLO MODELS FOR TEST DATA

Model	$mAP@0.5$	$mAP@0.75$	$mAP@[0.5:0.95]$
	Yolov8s	78.9	65.3
Yolov8m	81.2	68.7	53.8
Yolov8l	84.7	72.8	57.4
Yolov11s	79.5	66.9	51.2
Yolov11m	83.4	70.5	55.1
Yolov11l	86.9	75.3	59.2

In the analyses, YOLOv11 models demonstrated higher accuracy rates compared to YOLOv8 at the same scales. The superiority of YOLOv11 is particularly evident in narrower thresholds such as $mAP@0.75$ and $mAP@[0.5:0.95]$. YOLOv11 appears to be a more advanced update over YOLOv8. Notably, the large-scale YOLOv11l model stands out for its high accuracy in complex object detection scenarios.

TABLE IV. YOLO CLASS PERFORMANCE TABLE VALID DATA

Class	Images	Box(P)	R	$mAP@50$	$mAP@[0.5:0.95]$
Alzheimer	34	0.834	1.0	0.897	0.897
CONTROL	32	0.692	1.0	0.96	0.96
Parkinson	34	0.825	0.825	0.854	0.854

Throughout all epochs, the YOLOv11 model exhibited lower training loss compared to YOLOv8. This demonstrates that YOLOv11 has a faster learning capacity and a more effective optimization process. While YOLOv8 initially has higher training loss, it gradually decreases over successive epochs, stabilizing the model. In contrast, YOLOv11 consistently achieves lower validation loss than YOLOv8 across all epochs, indicating superior generalization on the validation dataset. However, YOLOv8 shows a less consistent decline in validation loss, reflecting a more irregular trend during its training process.

These results indicate that the YOLOv11 model is superior to the YOLOv8 model in both learning capacity and generalization ability. YOLOv11's lower training and validation losses demonstrate that the model is better optimized and has a lower risk of overfitting compared to YOLOv8. While YOLOv8 may be suitable for less complex datasets or scenarios with limited hardware resources, YOLOv11 is better suited for tasks requiring higher accuracy and generalization.

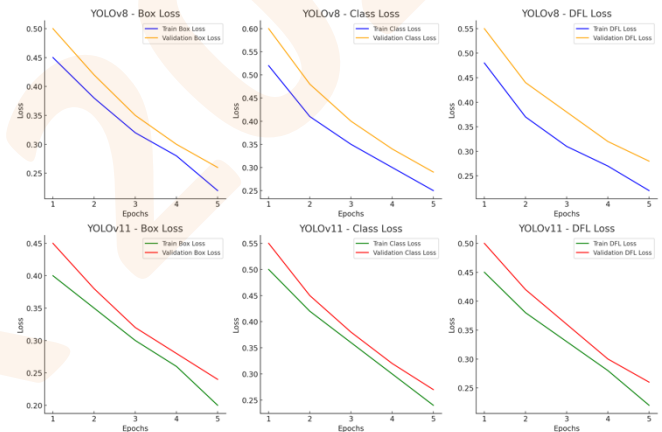


Fig. 5. Training and Validation Losses for YOLOv8 and YOLOv11

Figure 5 provides a detailed visualization of the training processes and validation performance of YOLOv8 and YOLOv11 models. YOLOv11 demonstrates superior performance in generalization, particularly by achieving lower losses in box loss and DFL loss metrics on validation data. However, YOLOv8 stands out for its rapid improvement during the training process.

These results suggest that YOLOv11 is better suited for more complex or focused distribution tasks, while YOLOv8 offers advantages in scenarios requiring quick convergence during shorter training cycles. The graphs serve as a valuable foundation for comparing the suitability of both models for different data types or application scenarios.

IV. CONCLUSION

In this study, YOLO models were utilized for the diagnosis of neurological diseases (Alzheimer, Parkinson, and control), deliberately eliminating the need for preprocessing. A total of 999 MRI images belonging to three different categories were used, with no preprocessing applied to the images. Experimental findings indicate that the YOLOv11 model demonstrates superior performance in tasks requiring complex or focused distributions, achieving low loss values and high accuracy rates. On the other hand, the YOLOv8 model offers advantages in shorter training periods due to its faster convergence capability. The results highlight that both models can be optimized for specific application scenarios and underscore the significant potential of deep learning-based methods in the diagnosis of neurological diseases.

REFERENCES

- [1] Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *The Lancet*, 386(9996), 896-912.
- [2] Wang, X., Zheng, W., Xie, J., & Wang, T. (2019). Neuroinflammation-mediated microglial activation in Alzheimer's disease and Parkinson's disease. *Progress in Neurobiology*, 179, 1-19.
- [3] Dementia. World Health Organization (WHO). Available from: <https://www.who.int/news-room/fact-sheets/detail/dementia>.
- [4] Smith, J. et al., "Neurodegenerative Diseases and Motor Function: An Overview," *Journal of Neurology*, 2020.
- [5] Aydın, S., Taşyürek, M., & Öztürk, C. (2022, May). MR Görüntüleri Ön İşlenerek Derin Ağlar ile Alzheimer Hastalığı Tespiti. In *International Conference on Emerging Sources in Science* (pp. 150-156).
- [6] Doe, A. et al., "The Role of MRI in Detecting Neurodegeneration in Parkinson's Disease," *Imaging in Medicine*, 2021.
- [7] Reddy, S., Giri, D., & Patel, R. (2024). Artificial Intelligence Diagnosis of Parkinson's Disease From MRI Scans. *Cureus*, 16(4).
- [8] Desai, S., Patel, K., Patel, A., Jadav, N. K., & Tanwar, S. (2024). A Survey on AI-based Parkinson Disease Detection: Taxonomy, Case Study, and Research Challenges. *Scalable Computing: Practice and Experience*, 25(3), 1402-1423.
- [9] Salehi, A. W., Baglat, P., Sharma, B. B., Gupta, G., & Upadhyaya, A. (2020, September). A CNN model: earlier diagnosis and classification of Alzheimer disease using MRI. In *2020 International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 156-161). IEEE.
- [10] Abd-Aljabar, S. S., Basheer, N. M., & Alsaif, O. I. (2022). Alzheimer's diseases classification using YOLOv2 object detection technique. *Int J Reconfigurable & Embedded Syst* ISSN, 2089(4864), 4864.
- [11] Islam, J., Furqon, E. N., Farady, I., Lung, C. W., & Lin, C. Y., "Early Alzheimer's Disease Detection Through YOLO-Based Detection of Hippocampus Region in MRI Images," In *2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)*, pp. 32-35, IEEE, 2023.
- [12] Alshahrani, A., Mustafa, J., Almatrafi, M., Albaqami, L., Aljabri, R., & Almontashri, S. (2024). A Comparative Study of Deep Learning Techniques for Alzheimer's disease Detection in Medical Radiography. *IJCSNS*, 24(5), 53.
- [13] Karthikeyan, B., Futina, J. J., Abinaya, P. P., & Priya, R. S. (2023, March). Parkinson's Disease Detection And Classification Of Stages From Drawing Patterns Using Deep Learning. In *2023 International Conference on Bio Signals, Images, and Instrumentation (ICBSII)* (pp. 1-7). IEEE.
- [14] Li, H., Yang, Z., Qi, W., Yu, X., Wu, J., & Li, H. (2024). Parkinson's image detection and classification based on deep learning. *BMC Medical Imaging*, 24(1), 187.
- [15] Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo Algorithm Developments. *ScienceDirect*. Retrieved from <https://doi.org/10.1016/j.procs.2022.01.135>
- [16] Alshahrani, A., Mustafa, J., Almatrafi, M., Albaqami, L., Aljabri, R., & Almontashri, S. (2024). A Comparative Study of Deep Learning Techniques for Alzheimer's disease Detection in Medical Radiography. *IJCSNS*, 24(5), 53.
- [17] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A., "You only look once: Unified, real-time object detection," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016.
- [18] Rafael Padilla, Wesley L Passos, Thadeu LB Dias, Sergio L Netto, and Eduardo AB Da Silva. A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics*, 10(3):279, 2021.
- [19] Akhilesh Sharma, Vipin Kumar, Louis Longchamps, "YOLO Algorithms for Object Detection: Comprehensive Survey and Future Perspectives," *Procedia Computer Science*, vol. 223, pp. 245-254, 2024. Available: <https://doi.org/10.1016/j.procs.2024.01.135>.
- [20] Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors. *arXiv preprint arXiv:2411.00201*.
- [21] Rahima Khanam* and Muhammad Hussain, YOLOv11: Enhancing Object Detection in Medical Imaging Through Programmable Gradient Features. *arXiv preprint*. <https://arxiv.org/html/2410.17725v1>
- [22] Mücaded Al Rabbani Elif, "YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications in Intelligent Transportation Systems," *ResearchGate preprint*, 2024. Available: https://www.researchgate.net/publication/385386644_YOLOv11_for_Vehicle_Detection_Advancements_Performance_and_Applications_in_Intelligent_Transportation_Systems.
- [23] Qin, X., Yu, C., Liu, B., & Zhang, Z. (2024). YOLO8-FASG: A High-Accuracy Fish Identification Method for Underwater Robotic System. *IEEE Access*.

A Business Analysis Model for Open Source Communities: Enhancing Collaboration and Innovation

Doğan Şengül
Department of Industrial Engineering,
Faculty of Engineering and Natural
Sciences
İstanbul Sabahattin Zaim University
İstanbul, Türkiye
0000-0002-2285-3907

Abstract— Open source communities play an essential role in fostering innovation and collaborative development. They operate in knowledge-based ecosystems characterized by transparency, inclusivity and a shared purpose. However, open source communities have faced challenges because of inefficient communication, insufficient knowledge sharing and poor requirements elicitation, management, prioritization and traceability. In this regard, this study proposes a business analysis model that is designed to address these challenges and to empower open source communities to succeed in greater collaboration and innovation as a result. Our proposed model aims to leverage key knowledge areas, tasks and elements of business analysis, aligned with the foundational values of open source, which are upstream first, community, and freedom. It emphasizes the components of (a) stakeholder-centric collaboration, (b) integrated knowledge management, (c) transparent decision-making, (d) freedom, licensing compliance and (e) iterative innovation and feedback cycles. Our model suggests significant contributions to both literature and practitioners by bridging the gap between business analysis knowledge area domains and the ecosystem mindset of the open source. It integrates structured techniques such as stakeholder mapping, documentation frameworks, and traceability practices into the open source context, addressing the critical challenges of fragmentation and inefficiency. The proposed model can have broad applications across industries such as software development, telecommunications, and in any branches of various sectors where open source collaboration drives efforts for advancements. By advocating shared vision, operational clarity and fair engagement, this study’s approach empowers communities to achieve scalable, innovative solutions while preserving the principles of openness and inclusivity.

Keywords — open source, business analysis, project management, requirements, collaboration, innovation

I. INTRODUCTION

Open source communities have become essential drivers of innovation in the modern digital landscape, revolutionizing how software is designed, developed, and maintained. Rooted in principles of inclusivity, transparency, and collective effort, these communities enable diverse contributors to collaborate toward common goals. As highlighted in the 2024 State of Open Source Report, over 95% of organizations have adopted open source software (OSS) attracted by its cost-effectiveness and adaptability [1, 2]. Moreover, open source platforms serve as critical and innovative hubs, facilitating both large-scale software programs with projects and niche community-driven initiatives [3, 4].

Effective project management and requirements engineering is a cornerstone of successful open source initiatives and projects. These processes ensure that the development efforts align with community goals and objectives, resources are efficiently utilized with respect to business drivers, and contributors and all stakeholders remain motivated and fulfilled. Recent research highlights that structured management and governance frameworks enhance collaboration and improve project outcomes by providing clear analysis and guidelines, roles, and decision-making protocols that are adaptable, flexible, and speedy [5, 6]. In addition, ensuring analysis for traceability and prioritization of requirements helps maintain focus, nimbleness and agility within the decentralized nature of OSS communities [7].

Having stated their many strengths, open source communities have faced several persistent challenges. The following section explores these challenges in detail and introduces a proposed model to address them.

II. CHALLENGES

Despite their achievements and contributions, open source communities face significant challenges that limit their full potential:

A. Inefficient Communication

The decentralized and concurrent communication channels prevalent in open source projects often lead to misunderstandings and coordination breakdowns. Studies demonstrate that while communication networks help enhance project performance to a great extent, overly busy networks can hinder productivity due to increased coordination overhead and redundant interactions [7] that overwhelm the communities. This relationship highlights the need for balance in communication concentration within these communities.

B. Insufficient Knowledge Sharing

Knowledge sharing is an important challenge in open source communities. The lack of centralized repositories and standardized documentation reduces the efficiency of knowledge, requirements and design reuse and onboarding of new contributors [4]. This fragmentation slows progress and also contributes to high attrition rates among less experienced participants.

C. Poor Requirements Elicitation and Traceability

Effective requirements management is vital to align stakeholders with project goals. However, open source projects often lack formalized approaches to eliciting, prioritizing, and tracking requirements. As [8, 9] note, the absence of structured management practices in many projects leads to fragmented development efforts and uncoordinated contributions. Furthermore, governance frameworks in some projects fail to effectively integrate diverse and beneficial stakeholder inputs, limiting the inclusivity and the efficiency of decision-making processes [5, 6].

D. Coordination and Productivity Issues

Larger team sizes in open source projects often lead to decreased productivity due to the Ringelmann effect, where increased coordination burden may reduce individual output [8]. This phenomenon highlights that the importance of adopting scaled coordination means to maintain efficiency as projects grow more.

E. Governance and Control:

Governance structures in open source communities vary widely, ranging from authoritarian to community-driven models. These variations can create inconsistencies in how projects handle governance, dealing with change-requests, information and communication management, decision-making, issue and conflict resolution. Effective governance mechanisms are essential to balance autonomy with accountability and responsibility, fostering an environment that motivates stakeholders while ensuring alignment with project goals and objectives [8, 9, 10].

Addressing these challenges requires the adoption of structured business analysis practices tailored to the very unique dynamics and synergies of open source communities. By enhancing communication strategies, developing requirements management, and implementing effective governance and frameworks, these very important communities can overcome barriers to collaboration and innovation.

III. BUSINESS ANALYSIS MODEL FOR OPEN SOURCE COMMUNITIES

Open source communities have redefined software development with their principles of collaboration, transparency, and inclusivity. However, achieving sustainable and efficient innovation in such decentralized ecosystems requires structured business analysis frameworks.

As stated in Business Analysis Body of Knowledge (BABOK) [11] business analysis is the process of identifying needs, defining requirements, recommending designs and solutions to deliver value to stakeholders. It involves the identification, assessment, and prioritization of requirements, as well as the development of strategies to achieve organizational objectives. According to [11], business analysis spans a variety of activities, including elicitation, documentation, verification, validation, traceability and stakeholder engagement, and plays a crucial role in aligning organizational efforts and stakeholder viewpoints with goals and objectives.

Building on the guidelines provided by BABOK and insights from recent studies, we present a comprehensive business analysis model designed to address the very inherent challenges of open source collaboration. Building from the

Business Analysis Core Concept Model of BABOK, this study presents the key components of the proposed model as follows, as demonstrated in Fig. 1:

A. Stakeholder-Centric Collaboration

Effective collaboration in open source projects necessitates active stakeholder engagement. BABOK highlights that identifying and managing diverse stakeholders is crucial for aligning objectives and fostering participation [11]. Research shows that collaboration dynamics in open source platforms heavily rely on structured interactions among core contributors and participants [4]. For instance, the "onion model" categorizes the discovery of contributors by their level of commitment, suggesting strategies for integrating new stakeholders while maintaining productivity and accountability [11, 12]. The model incorporates stakeholder mapping and traceability tools to identify the roles and interests of stakeholders, namely contributors, maintainers, and end-users. Additionally, transparent communication means are established to enable equitable participation and collective decision-making.

Several business analysis techniques can be effectively applied to generate and enhance stakeholder-centric collaboration in open source projects. Stakeholder analysis helps identify and categorize stakeholders based on their influence, interests, and levels of engagement. Techniques like RACI and stakeholder matrices provide clarity on roles and responsibilities, ensuring alignment with project objectives [11]. Workshops and focus groups enable effective communication among stakeholders, developing collaboration and gathering diverse perspectives. Additionally, elicitation techniques, such as interviews and surveys, help gather insights into stakeholder needs, enabling more informed decision-making and engagement strategies.

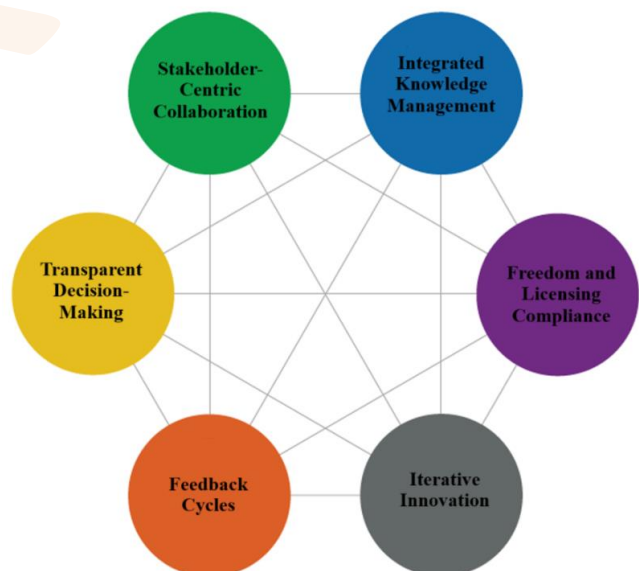


Fig. 1. Key components of the proposed model

B. Integrated Knowledge Management

Fragmented knowledge repositories are a common issue in open source projects, leading to inefficiencies in onboarding and collaboration. The importance of centralized knowledge management is emphasized in [11] as elicitation and collaboration guidelines, which advocate for reusable existing documentation. Studies highlight the impact of well-

maintained repositories on improving project responsiveness and reducing the learning curve for new contributors [13].

To address the challenges of fragmented knowledge repositories in open source projects, several business analysis techniques from BABOK can be employed [11]. Document analysis ensures the reuse and organization of existing knowledge assets, streamlining onboarding and collaboration. Traceability management tracks the alignment of documentation with project goals, enabling contributors to identify dependencies and gaps effectively [11]. Metrics and indicators can be used to evaluate the effectiveness of knowledge management practices, ensuring continuous improvement [13].

C. Transparent Decision-Making

Governance frameworks are critical for maintaining transparency and accountability in open source ecosystems. BABOK emphasizes planning and monitoring as key components of effective decision-making processes [11]. Research has identified multiple governance models, including defined community and open community configurations, which combine varying degrees of formal and informal controls to effectively manage contributions [5].

The model integrates governance mechanisms such as voting systems and prioritization frameworks to ensure transparency in decision-making. These tools not only foster trust among the contributors but also streamline the approval and integration of new features and bug fixes.

To ensure transparency and accountability in decision-making, several business analysis techniques from BABOK can be applied effectively. Decision analysis provides a structured approach to evaluating and prioritizing options, enabling communities to make fair and informed decisions. Governance frameworks define clear roles, responsibilities, and processes for handling conflicts and resource allocation. Workshops and collaborative sessions encourage stakeholder participation in prioritization and consensus-building activities. Also, prioritization techniques such as MoSCoW are instrumental in managing competing demands for limited resources, ensuring alignment with project goals. Root cause analysis can also be employed to address decision-making bottlenecks, improving the overall governance process [11].

D. Freedom and Licensing Compliance

Adhering to open source licensing principles is essential for preserving community values and ensuring legal compliance. Studies reveal that restrictive licenses may discourage developer participation while encouraging end-user adoption, underscoring the need for balanced licensing strategies [12, 14]. BABOK's compliance management guidelines provide a structured approach to navigating these complexities [11]. This model includes tools for tracking licensing obligations and promoting awareness among contributors about compliance requirements. Clear guidelines for contributions and code reuse help maintain community integrity and trust.

Business analysis techniques like document analysis, reviews and item tracking provide structured methods to track licensing obligations and ensure adherence to open source principles. Risk Analysis helps identify potential legal risks, while stakeholder collaboration ensures contributors understand compliance requirements. These techniques

collectively promote transparency and safeguard community integrity [12].

E. Iterative Innovation and Feedback Cycles

Continuous improvement is at the heart of open source development. Agile-inspired iterative methods, supported by BABOK's solution evaluation framework, ensure that projects remain responsive to user needs [11]. Empirical studies demonstrate that regular feedback cycles enhance problem-solving capabilities and foster innovation in dynamic team structures [7]. The proposed model incorporates iterative planning sessions, retrospective reviews, and real-time feedback mechanisms. These practices enable teams to address challenges promptly and adapt to evolving requirements effectively.

Business analysis techniques like backlog management, retrospective reviews and solution evaluation knowledge area from BABOK support iterative innovation by enabling communities to assess progress and incorporate feedback continuously. Prototyping and metrics further ensure responsiveness to evolving requirements while fostering adaptability and problem-solving efficiency [7, 11].

By integrating structured business analysis practices with open source principles, the model we present addresses key challenges such as stakeholder misalignment, fragmented knowledge, and governance gaps. Leveraging insights from BABOK and challenges put forward by recent research, it provides a robust framework for enhancing collaboration and innovation in open source communities. Ultimately, this approach empowers contributors and stakeholders as a whole to achieve shared goals while preserving the ethos and mottos of openness and inclusivity.

IV. METHODOLOGICAL FRAMEWORK OF THE PROPOSED MODEL

The methodological framework for the proposed business analysis model, as demonstrated in Fig. 2, begins with identifying the business case, key goals, challenges and conducting stakeholder analysis to map roles and responsibilities. The framework then incorporates collaboration mechanisms for participation and transparent decision-making. To streamline knowledge sharing, the framework emphasizes centralized knowledge management, where repositories and traceability tools align documentation with project objectives. Governance is strengthened through structured frameworks and decision-making tools like prioritization techniques. Licensing compliance is integrated through risk analysis, in order to be in line with open source principles while maintaining community trust. The framework also adopts iterative development practices inspired by adaptive approaches, including backlog management and feedback loops, to adapt to evolving requirements and foster innovation. Monitoring with performance metrics and related KPIs will ensure that the framework remains aligned with community needs, promoting sustainable collaboration, scalability, and inclusivity. This structured and flexible approach aims to bridge formal business analysis practices and the organic operations of open source communities.



Fig. 2. Main flow of the methodological framework of the proposed model

V. APPLICATIONS AND IMPACT

The proposed business analysis model has significant and broad applications across different industries where open source collaboration serves as a driver for innovation. Software development projects benefit from streamlined analysis, governance and enhanced knowledge sharing, while sectors like telecommunications may leverage iterative development to integrate cutting-edge technologies and adapt to changing market demands.

A. Contributions to Literature and Practitioners

This model bridges between structured business analysis practices and the decentralized nature of open source communities and ecosystems. By incorporating stakeholder-centric collaboration, integrated knowledge management, and adaptive feedback loops, the model addresses key challenges such as fragmented governance and inefficient communication [10]. It is emphasized that the competitive advantage is gained by projects that prioritize user responsiveness and efficient defect-fixing processes [15, 16]. The model also extends existing research by introducing business analysis techniques like governance and prioritization frameworks to the open source context, offering practicality for resolving inefficiencies.

B. Future Directions

While the model provides a robust framework, further research could explore its application to hybrid open source-commercial projects, where additional issues arise from differing stakeholder priorities. Moreover, integrating predictive analytics in governance and collaboration frameworks offers promising research areas for enhancing project efficiency in terms of budget, scope, quality and schedule.

Moreover, future studies could assess the long-term impact of iterative feedback cycles on project performance across different cultural and organizational contexts. This exploration could provide valuable insights into tailoring the model for implementation across different industries and initiatives, thereby making it more scalable and adaptable to varying environments.

REFERENCES

- [1] Perforce Software, 2024 State of Open Source Report: Open Source Usage, Market Trends, & Analysis, Perforce Software, 2024. [Online]. Available: <https://www.openlogic.com/sites/default/files/pdfs/report-ol-state-of-oss-2024.pdf>
- [2] Eclipse Foundation, *How Open Source Working Groups Drive Collaboration and Innovation*, Eclipse Foundation, 2024. [Online]. Available: <https://outreach.eclipse.foundation/working-groups-open-source-ebook>
- [3] E. S. Raymond, *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, Revised ed. Sebastopol, CA: O'Reilly Media, 1999.
- [4] N. Zöller, J. H. Morgan, and T. Schröder, "A topology of groups: What GitHub can tell us about online collaboration," *Technological Forecasting and Social Change*, vol. 161, p. 120291, 2020. [Online]. Available: <https://doi.org/10.1016/j.techfore.2020.120291>
- [5] D. Di Tullio and D. S. Staples, "The governance and control of open source software projects," *Journal of Management Information Systems*, vol. 30, no. 3, pp. 49–80, Dec. 2014. [Online]. Available: <https://doi.org/10.2753/MIS0742-1222300303>.
- [6] M. L. Markus, "The governance of free/open source software projects: Monolithic, multidimensional, or configurational?," *Journal of Management Governance*, vol. 11, pp. 151–163, 2007. [Online]. Available: <https://doi.org/10.1007/s10997-007-9021-x>.
- [7] S. Pal, A. Nair, and Z. Zuo, "Collaborative dynamics in open source software development: Unveiling the influence of team interaction and the role of project manager," *Journal of Operations Management*, vol. 70, no. 7, pp. 1076–1099, 2024. [Online]. Available: <https://doi.org/10.1002/joom.1324>.
- [8] I. Scholtes, P. Mavrodiev, and F. Schweitzer, "From Aristotle to Ringelmann: A large-scale analysis of team productivity and coordination in open source software," *Empirical Software Engineering*, vol. 21, no. 2, pp. 642–687, 2016. [Online]. Available: <https://doi.org/10.1007/s10664-015-9406-4>.
- [9] J. Hahn and C. Zhang, "An exploratory study of open source projects from a project management perspective," *Krannert School of Management, Purdue University*, 2005. [Online]. Available: http://www.krannert.purdue.edu/academics/mis/workshop/hz_110405.pdf. Accessed: Oct. 22, 2024.
- [10] A. H. Ghapanchi, C. Wohlin, and A. Aurum, "Resources contributing to gaining competitive advantage for open source software projects: An application of resource-based theory," *International Journal of Project Management*, vol. 32, no. 1, pp. 139–152, 2014. [Online]. Available: <https://doi.org/10.1016/j.ijproman.2013.03.002>.
- [11] IIBA, *A Guide to the Business Analysis Body of Knowledge (BABOK Guide)*, v3. Toronto, ON, Canada: International Institute of Business Analysis, 2015.
- [12] C. Connell, "Open source projects manage themselves? Dream on," *IBM/Lotus Developers Network*, 2000. [Online]. Available: http://www.chc-3.com/pub/manage_themselves.htm. Accessed: Nov. 21, 2024.
- [13] J. F. Schrape, "Open source communities: The sociotechnical institutionalization of collective invention," in *Collectivity and Power on the Internet*, SpringerBriefs in Sociology. Cham, Switzerland: Springer, 2018. [Online]. Available: https://doi.org/10.1007/978-3-319-78414-4_4
- [14] I. Stamelos, "Management and coordination of free/open source projects," in *Software Project Management in a Changing World*, G. Ruhe and C. Wohlin, Eds. New York: Springer, 2014, pp. 321–341.
- [15] C. Subramaniam, R. Sen, and M. L. Nelson, "Determinants of open source software project success: A longitudinal study," *Decision Support Systems*, vol. 46, no. 2, pp. 576–585, 2009.
- [16] R. Ng, *Free and Open-Source Software: Coordination and Competition*. Collaborative Research Center Transregio 224, Discussion Paper No. 585, University of Mannheim, 2024. [Online]. Available: https://crctr224.de/discussion_papers/Ng_FOSS_Coordination_Competition. Accessed: Nov. 21, 2024.

Açık Kaynak Kodlu Araçlar ile RGB Kamera Kullanarak Gezgin Robot Projesi Oluşturma

Creating a Mobile Robot Project Using RGB Camera with Open Source Tools

Kadir Aram
Bilgisayar Mühendisliği
Fatih Sultan Mehmet Vakıf Üniversitesi
İstanbul Türkiye
karam@fsm.edu.tr

Rami Eyüboğlu
Bilgisayar Mühendisliği
Fatih Sultan Mehmet Vakıf Üniversitesi
İstanbul Türkiye
rami.ayoub@stu.fsm.edu.tr

Ahmad Karam Ahmedoğlu
Bilgisayar Mühendisliği
Fatih Sultan Mehmet Vakıf Üniversitesi
İstanbul Türkiye
ahmadkaram.alahmad@stu.fsm.edu.tr

Özetçe— Robotlar, diğer teknolojik cihazlar gibi, insan hayatına fayda sağlamak amacıyla birçok alanda kullanılmaktadır. Özellikle doğal afetler ve askeri uygulamalar gibi kritik alanlarda robotların kullanımı giderek daha önemli hale gelmiştir. Bu süreç, robotların geliştirilmesi ve belirli görevleri başarıyla yerine getirmelerinin sağlanmasıyla başlamaktadır. Robotik sistemlerin geliştirilmesi sırasında pek çok platform ve yazılım çerçevesi kullanılmakta olup, bu çalışmada açık kaynak kodlu platform ve yazılımlar ele alınmıştır. Çalışmada, gezgin robotlar ile proje oluşturma sürecinde bu açık kaynak sistemlerin kullanımını uygulama örnekleriyle açıklamak amacıyla hem kara hem de hava robotları incelenmiştir. Deneysel çalışmalar sırasında, açık kaynak sistemler kullanılarak bir robotun engellerle dolu bir ortamda başarılı bir şekilde navigasyon gerçekleştirdiği gösterilmiştir. Robot, RGB kamera ile çevresinden aldığı görüntülerle engelleri tespit etmiş ve hedef noktaya ulaşmıştır. Çalışmada robot olarak açık kaynak yazılımlarını destekleyen TurtleBot kullanılmış, simülasyon ortamı olarak ise Gazebo tercih edilmiştir. Bu araştırma, açık kaynak kodlu platformların gezgin robot uygulamalarındaki rolüne dair önemli bulgular sunmaktadır.

Anahtar Kelimeler — Açık kaynak; gezgin robot; görüntü işleme;kamera.

Abstract— Robots, like other technological devices, are increasingly utilized across various fields to enhance human life. Their application is particularly significant in critical areas such as natural disasters and military operations. The development of robotic systems involves designing robots and assigning them specific tasks to fulfill autonomously. This process necessitates the use of numerous platforms and software frameworks. In this study, open-source platforms and software are explored, focusing on their role in mobile robot project development. To illustrate their utility, both ground and aerial robots were examined. Experimental results demonstrated the successful navigation of a robot in an obstacle-filled environment using open-source systems. The robot utilized RGB camera data to detect obstacles and reach the target location. TurtleBot, which supports open-source software, was employed as the robot, and the Gazebo simulation environment was used for testing. This research highlights the critical role of open-source platforms in mobile robotic applications and their adaptability for diverse scenarios.

Keywords — open source; mobile robot; image processing; camera.

I. GİRİŞ

İnsansız sistemler, temel veri işleme bileşenleri, sensörler, otomatik kontrol mekanizmaları ve iletişim sistemleriyle donatılmış, bağımsız görevler gerçekleştirebilen mekanik yapılar veya cihazlardır. Bu tür sistemler, insansız hava araçları (İHA), insansız kara araçları (İKA) ve su altı keşif cihazları gibi geniş bir yelpazeye yayılan araçları kapsar. İnsansız sistemler, otonom hareket kabiliyetleri sayesinde keşif, gözetim, haritalama ve analiz gibi çeşitli görevleri güvenilir bir şekilde yerine getirebilirler [1]. Gezgin robotlar ise, belirli bir görev veya hedef doğrultusunda, tanımlanmış bir çalışma alanı içinde özerk olarak hareket edebilen robotlardır. Bu robotlar, insan müdahalesine ihtiyaç duymaksızın belirlenen hedeflere ulaşmak amacıyla serbestçe yön değiştirebilme ve engellerden kaçınma yeteneğine sahiptir. Robotların bu üstün özellikleri, tarım, kargo taşımacılığı, haritalama, arama-kurtarma gibi çok çeşitli uygulama alanlarında hedeflerine etkin bir biçimde ulaşmalarını mümkün kılmıştır [2].

Açık kaynaklı sistemler(AKS), kullanıcılara kaynak kodunu herhangi bir amaçla çalıştırma, inceleme, değiştirme ve özgün veya değiştirilmiş biçimde yeniden dağıtma hakları veren bir lisans altında yayınlanan herhangi bir sistemdir. Bu özelliği sayesinde toplulukların katkısına imkân tanımaktadır. Günümüzde açık kaynak kodlu yazılımlar, bilimsel araştırmalardan ticari uygulamalara kadar birçok alanda kullanılmaktadır. Bu sistemler, kullanıcılardan geri bildirimler alarak, sürekli iyileştirme, hızlı bir şekilde güncelleme ve uygulama esnekliği sağlarlar[3], [4].

Açık kaynak kodlu sistemlerin bu özellikleri onları robotik alanında da sık kullanılmasını sağlamıştır. Özellikle otonom hareket, engelden kaçınma, haritalama ve konum belirleme gibi temel mobil robot işlevleri için açık kaynak araç ve çerçeveler, araştırmacılar ve mühendisler tarafından tercih edilmektedir. Bu tür araçlar, araştırmacılara algoritmaları daha fazla kontrol etmeleri sayesinde ve yenilikçi çözümler geliştirmelerini sağlamaktadırlar. Açık kaynak mobil robot platformları, robotik araştırma ve geliştirme süreçlerini hızlandırmak için çok sayıda araç sunmaktadır. Robot Operating System(ROS) gibi platformlar, mobil robotlar için güçlü bir yazılım altyapısı sunarak sensör entegrasyonu, hareket planlaması ve otonom kontrol gibi alanlarda geniş bir uygulama yelpazesi

sunmaktadır [5]. Açık kaynaklı simülasyon ortamları, mobil robotların gerçek dünyadaki performanslarını simüle etmek için kullanılır ve araştırma maliyetlerini azaltır. Bu simülatörler, robotik algoritmaların hızlı bir şekilde test edilmesi ve optimize edilmesi için olanak sağlar ve fiziksel ortam gereksinimlerini en aza indirir [6].

Yapılan çalışmalar açık tamamen açık kaynak kod bir sistem üretmek ya da açık kaynak kod platformları kullanarak yeni projeler geliştirme olarak ayrıştırılabilir. Lopez ve çalışma arkadaşları, açık kaynaklı, düşük maliyetli, modüler ve genişletilebilir bir mobil robot geliştirmiştir. Bu robotun tasarımında, açık kaynaklı Android ve Arduino tabanlı cihazlar ile yazılımlar kullanılmıştır. Geliştirilen mobil robot, özellikle sınıf içi veya laboratuvar ortamları için geleneksel görsel laboratuvarlara bir alternatif olarak tasarlanmış olup, uzaktan eğitim ve geniş çaplı açık çevrimiçi dersler gibi modern eğitim uygulamalarında kullanılmak üzere optimize edilmiştir [7]. Başka bir çalışmada açık kaynak olarak sunulmuş, kapalı döngü ters kinematik kontrolü ile serbest hareket ve navigasyonu mümkün kılan sensör tabanlı bir mobil robotik mimari geliştirilmiştir. Bu tür robotlar, özellikle engellerin genellikle hareket ettiremediği ve dar koridorların bulunduğu kapalı endüstriyel ortamlarda, işbirlikçi kapalı alan senaryolarında kullanılmaktadır [8]. Kohlbrecher ve diğerleri, bozulmuş kentsel ortamlarda robotların kendi kendini konumlandırma ve haritalama (SLAM) yeteneklerini geliştirmek amacıyla Hector SLAM algoritmasını içeren açık kaynaklı yazılım modüllerini tanıtmıştır [9]. Bir diğer çalışmada açık kaynak platformlar ile hava ve karasal robotları içeren robot sürüsü oluşturulmuş ve bu robotların formasyon kontrolü ve navigasyonu gerçekleştirilmiştir [10].

Bu çalışmada, açık kaynak kodlu platformlar kullanılarak bir gezgin robot projesi oluşturulma süreçleri ele alınmıştır. Çalışma, otonom hareket kabiliyeti kazandırılmış bir gezgin robotun, çevresindeki engelleri algılayarak güvenli bir şekilde hedefe ulaşmasını sağlamayı amaçlamaktadır. Robot, bir RGB kamera ile çevresel verileri toplamakta ve bu verileri gerçek zamanlı olarak işleyerek engeli tespit etmekte ve hedef noktaya ulaşmaktadır. Çalışmanın ikinci kısmında robotik uygulamalarda kullanılan açık kaynak kodlu sistemlerden bahsedilmiş, üçüncü kısımda yapılan çalışma ve simülasyon sonuçları gösterilmiştir, dördüncü kısımda ise sonuçlar ve gelecek çalışmalar anlatılmıştır.

II. MATERYAL VE METOD

A. Materyaller

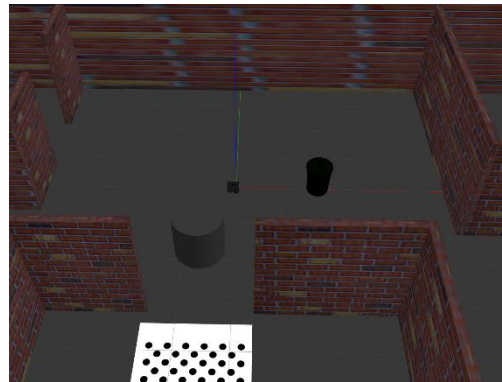
Bir mobil robot projesi geliştirme sürecinde, farklı aşamalarda birden fazla platformun etkili bir şekilde kullanımı gereklidir. Öncelikle, çözülmek istenen probleme uygun bir robotun ve çalışma ortamının seçilmesi kritik bir adımdır. Bu seçim, mobil robotun fiziksel özellikleri, kullanılacak sensörler ve hedef senaryoya uygun diğer donanım ve yazılım bileşenlerini içerir. Daha sonra, robotun görevini başarıyla yerine getirebilmesi için gerekli algoritmalar belirlenir ve bu algoritmaların test edilip geliştirilmesi amacıyla uygun bir simülasyon ortamı sağlanır.

Robot Operating System (ROS), ticari ve araştırma tabanlı robotik uygulamalarda yaygın olarak kullanılan, açık kaynaklı bir yazılım geliştirme ortamıdır [11]. Çok yönlülüğü

sayesinde, ROS robotik projelerinde sıklıkla tercih edilmektedir. Bu esneklik, farklı programlama dillerinde geliştirme imkânı sunması, yayıncı-abone mimarisi üzerinden süreçler arasında iletişim sağlaması, projeye özgü mesaj türlerini desteklemesi ve uygulamaların kolayca paketlenip paylaşılmasına olanak tanimasından kaynaklanmaktadır. Aynı zamanda ROS, yaygın robotik problemleri çözmek için geniş bir paket kütüphanesi ve güçlü araç desteği sunar. Örneğin, veri görselleştirme için kullanılan RVIZ, ROS'un sağladığı önemli araçlardan biridir [12]. ROS sisteminde süreçler "düğüm" (node) olarak adlandırılır ve bu düğümler, belirli bir iletişim mimarisi üzerinden birbirleriyle veri alışverişi yapar. Yayıncı düğümler, belirli bir konuda (topic) mesaj gönderirken, abone düğümler bu mesajları alarak işleyebilir. Bu yapı, robotik projelerde veri akışının düzenlenmesini ve modüler bir tasarımın oluşturulmasını sağlar [13].

ROS2 ise, ROS'un sınırlamalarını gidermek amacıyla geliştirilmiş bir yazılım platformudur. Özellikle ROS'un merkezi ağ yapısı, ağ güvenliği mekanizmalarının eksikliği ve mikroişlemcilerle entegrasyonun yalnızca üçüncü taraf yazılımlara bağımlı olması gibi kısıtlamalar ROS 2 ile ele alınmıştır. ROS 2, dağıtılmış iletişim mimarisi kullanarak, sistemler arası daha yüksek güvenilirlik ve ölçeklenebilirlik sağlamaktadır [14].

Webots, hem akademik hem de endüstriyel alanlarda robotik araştırma ve geliştirme için kullanılan güçlü bir robot simülasyon platformudur. Bu ortam, kullanıcıların üç boyutlu ortamlar tasarlamasına ve bu ortamlarda simüle edilen robotları çalıştırmasına olanak tanır [15]. Gazebo, robotların iç ve dış ortamlarda çalıştırılmasını sağlayan 3B simülasyon platformudur. Robotik araştırma projelerinde, yarışmalarda ve ticari uygulamalarda sıklıkla kullanılır. Gazebo, ODE, Bullet, DART ve Simbody dahil olmak üzere birden fazla fizik motorunu destekler. Özellikle kullanışlı olan, robot modelleri ve çevre dünyalarının hazırlanmış örnekleriyle birlikte gelmesidir. Gazebo topluluğu yeni robot modelleri ve çevre dünyaları oluşturmak için aktiftir [6]. Örnek bir simülasyon görüntüsü Şekil 1'de gösterilmektedir.



Şekil 1: Gazebo Simülasyon Ortamı

Turtlebot hem araştırma hem de eğitimde kullanılan ROS tabanlı bir gezgin robottur. Çalışmada mobil robot olarak Turtlebot3 Waffle Pi modeli kullanılmıştır.

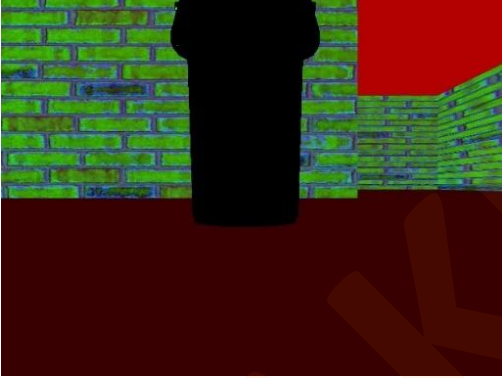
OpenCV (Open Source Computer Vision Library), açık kaynak kodlu bir bilgisayar görüşü ve makine öğrenimi yazılımıdır. Hem akademik hem de endüstriyel uygulamalarda yaygın olarak kullanılan bu kütüphane, bilgisayar görüşü alanında 2.500'den fazla optimize edilmiş

algoritma içermektedir. Bu algoritmalar, nesne tespiti, yüz algılama ve tanıma, insan hareket analizleri gibi uygulamalarda önemli bir performans sağlar. Ek olarak, stereo kameralar kullanılarak üç boyutlu (3B) nokta bulutlarının oluşturulması gibi ileri düzey bilgisayar görüşü işlemleri için de kullanılabilir [16].

B. Metod

Bu çalışmada, mobil robotun başlangıç noktasından hedef noktaya ulaşması sırasında engellerden sakınarak hareket etmesi sağlanmıştır. Bu süreç iki temel aşamada gerçekleştirilmiştir. İlk aşamada, robotun mevcut konumu ile hedef konum arasındaki navigasyon işlemi ele alınmıştır. İkinci aşamada ise, robotun hareket rotası üzerindeki engellerin tespit edilmesi ve bu engellerden etkin bir şekilde kaçınılması işlemleri gerçekleştirilmiştir.

1) *Engel tespiti*: Engel tespiti ve kaçınma aşamalarında, OpenCV kütüphanesi önemli bir araç olarak kullanılmıştır. Bu süreçte, robotun kamerasından elde edilen görüntüler analiz edilmiş ve belirli renk aralıklarına sahip bölgeler belirlenmiştir. Seçilen bu bölgeler üzerinde morfolojik işlemler uygulanarak, engellerin şekli ve sınırları başarıyla tanımlanmıştır. Bu yöntem sayesinde, robotun çevresel algısı geliştirilmiş ve hedefe güvenli bir şekilde ulaşması sağlanmıştır. Kameralardan elde edilen görüntünün HSV renk uzayındaki görünümü Şekil 2'de gösterilmiştir.



Şekil 2: Ortamın HSV Uzayındaki Görünümü

İlk adımda, elde edilen görüntüye aşındırma (erosion) işlemi uygulanmıştır. Bu işlem, görüntü üzerindeki gürültülü ve istenmeyen küçük parazit alanları ortadan kaldırmak için kullanılmıştır. Bu adım, nesnenin sınırlarını daha belirgin hale getirirken, çevresel parazitlerin robotun algılama performansını etkilemesini engeller. Bu yöntem, özellikle dış ortam koşullarında, karmaşık arka planların etkisini azaltarak engelin daha doğru bir şekilde tanımlanmasını sağlar. Şekil 3'te görüntünün aşındırma işleminden önce ve sonraki durumu gösterilmiştir.



Şekil 3: Sol tarafta Aşındırma Öncesi, Sağ tarafta Aşındırma Sonrası

Sıradaki aşamada, görüntü üzerinde bulunan gürültüleri azaltmak için bir dizi ön işleme adımı gerçekleştirilmiştir. Bu amaçla, görüntüye alçak geçiren bir filtre uygulanarak hafif bir bulanıklaştırma işlemi gerçekleştirilmiştir. Bu işlem, görüntüdeki yüksek frekanslı gürültü bileşenlerini ortadan kaldırarak, detayların bozulmasına yol açmadan gürültünün azaltılmasını sağlar. Ardından, görüntünün daha temiz ve işlenebilir bir hale gelmesi için "açılma" işlemi uygulanmıştır. Bu işlem, görüntü üzerindeki küçük parazitlerin giderilmesine ve nesne kenarlarının daha belirgin hale getirilmesine yardımcı olur. Şekil 4'te bu işlemler gösterilmiştir.



Şekil 4: Sol tarafta Bulanıklaştırma, Sağ tarafta Açılma İşlemi

Üçüncü aşamada, nesnenin sınırlarının ve şeklini ortaya çıkarmak için kontur bulma (contour detection) yöntemi kullanılmıştır. Bu işlemde, görüntüde aynı yoğunluğa sahip pikseller birleştirilerek bir eğri oluşturulur ve böylece nesnenin dış hatları tespit edilir. Kontur bulma işlemi, nesnenin geometrik özelliklerinin çıkarılmasını ve sahnedeki engellerin tanımlanmasını sağlar. Bu adım sonucunda elde edilen veriler, mobil robotun çevresindeki engelleri algılamasına ve bu engellere uygun şekilde hareket etmesine olanak tanır. Şekil 5'te işlem sonucu gösterilmiştir.



Şekil 5: Kontur Bulma İşlemi

Bu işlem dizisi, mobil robotların çevresel engelleri doğru bir şekilde algılaması ve bunlardan kaçınarak hedefe yönelmesi için kritik bir öneme sahiptir.

2) *Navigasyon*: Robotun başlangıç noktasından hedef noktaya ulaşabilmesi için uygun bir hareket planlaması yapılması gerekmektedir. Hedef nokta belirlendikten sonra, robot ile hedef arasındaki mesafe ve açı değerlerinin hesaplanması kritik bir öneme sahiptir. Bu hesaplamalar, robotun bulunduğu konumun sürekli olarak güncellenmesini ve hedefe doğru yönelmesini sağlar. Robotun hareketi sırasında mesafe ve yön bilgileri dinamik olarak yeniden değerlendirilerek, hedefe ulaşması için optimal bir yol izlenir.

Robotun başlangıç noktası ile hedef noktası arasındaki mesafe, Denklem 1'de hesaplanmıştır. Benzer şekilde, robotun hedefe yönelik açılma pozisyonu ise Denklem 2 ile belirlenmiştir. Robotun hareket sırasında güvenli ve kararlı

bir ilerleme sağlaması adına, hız parametreleri sınırlandırılmıştır. Bu bağlamda, robotun doğrusal hızı 0.2 m/s, açısal hızı ise 0.5 rad/s olarak tanımlanmıştır. Bu hız limitleri, robotun çevresel algısının doğru bir şekilde işlenmesi ve engellerden kaçınarak hedefe güvenli bir şekilde ulaşması için gereklidir.

$$mesafe(d) = \sqrt{(x_H - x_R)^2 + (y_H - y_R)^2} \quad (1)$$

$$\theta = \arctan\left(\frac{y_H - y_R}{x_H - x_R}\right) \quad (2)$$

Burada,

x_R, y_R : Robotun pozisyonu

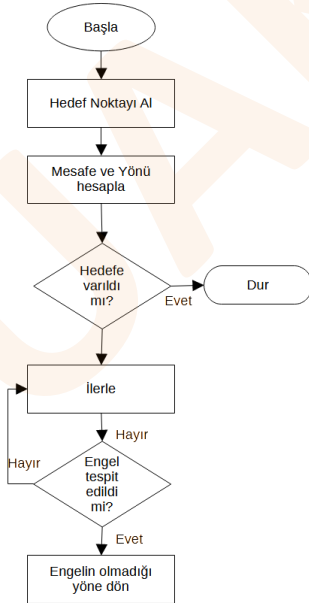
x_H, y_H : Hedef pozisyonu

θ : Robot ile hedef nokta arasındaki açıyı gösterir.

Bu işlemlerin tüm aşamaları, Şekil 6'daki akış diyagramında detaylı bir şekilde gösterilmiştir. İlk adımda, hedef noktanın tanımlanması işlemi gerçekleştirilmiştir. Hedef nokta belirlendikten sonra, robot ile hedef arasındaki mesafe ve açısal konum, Denklem (1) ve Denklem (2) kullanılarak hesaplanmıştır. Bu bilgiler doğrultusunda robot, yönünü belirleyerek hedefe ulaşmak için hareket etmeye başlar.

Robotun hareketi sırasında kameradan alınan görüntüler düzenli olarak analiz edilir. Eğer görüntü verilerinde bir engelin varlığı tespit edilirse, robot engelden kaçınmak için yönünü değiştirir. Engelin başarılı bir şekilde aşılmasının ardından, robot hedef noktaya doğru yeniden hareket etmeye devam eder. Bu süreç, robotun çevresel algısını ve hareket planlamasını dinamik bir şekilde gerçekleştirdiğini göstermektedir.

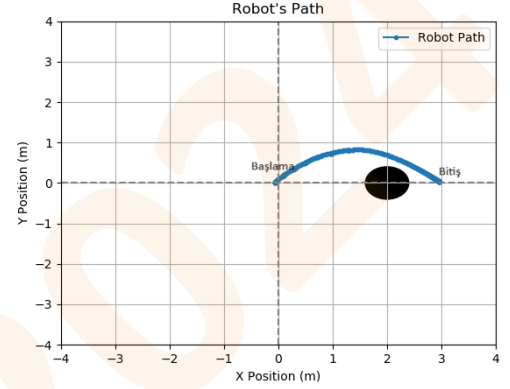
Bu yaklaşım, robotik sistemlerin gerçek zamanlı karar verme mekanizmalarının bir örneğini sunmaktadır. Hem hedefe ulaşma hem de engelden kaçınma görevlerini eşzamanlı olarak yerine getiren bu yapı, otonom navigasyon algoritmalarının etkinliğini ve pratik uygulamalardaki önemini vurgulamaktadır.



Şekil 6: Sistemin Akış Diyagramı

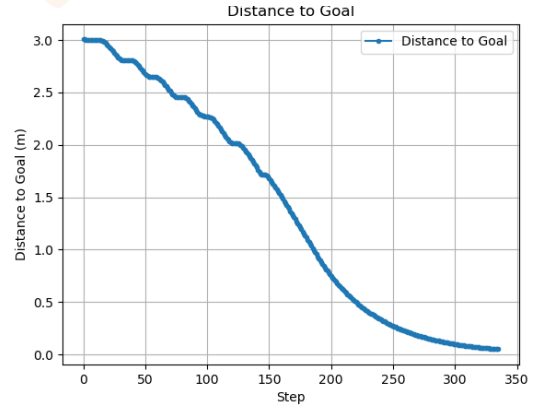
III. SIMÜLASYON VE TEST

Çalışma ROS Noetic sürümü kullanılarak Ubuntu üzerinde yürütülmüştür. Geliştirilen sistemin simülasyonu için iç ve dış mekanlarda kullanılabilen açık kaynak kodlu Gazebo programı kullanılmıştır. Simülasyonda; eğitim, araştırma ve ürün prototipleme vb. uygulamalarda sıklıkla kullanılan, programlanabilir, açık kaynak kodlu sistemleri destekleyen ROS tabanlı Turtlebot3 robotu kullanılmıştır. Oluşturulan yöntemlerin performansını değerlendirmek amacıyla, simülasyon farklı konum parametreleri ile çalıştırılmış ve sonuçlar detaylı bir şekilde gözlemlenmiştir. İlk deneyde, robot (0,0) başlangıç noktasından (0,3) hedef noktasına ilerlemiştir. Robotun bu süreçte izlediği yol Şekil 7'de sunulmuştur.



Şekil 7: Deney 1 Robotun yolu

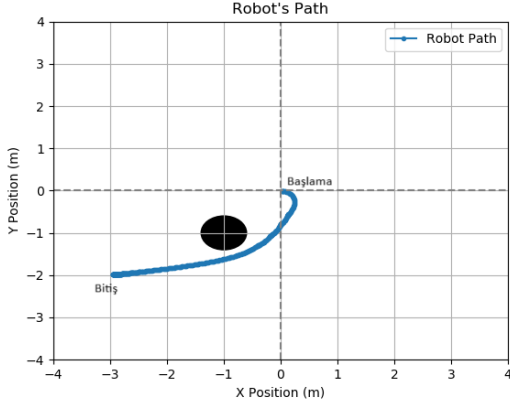
Deney sırasında robotun hedef noktaya olan uzaklığının grafiği de Şekil 8'de verilmiştir.



Şekil 8:Deney 1 Robotun hedef konuma uzaklığı

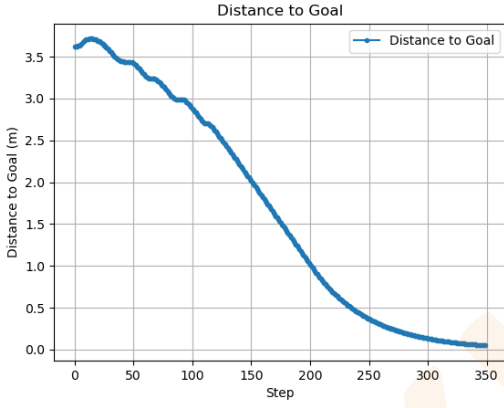
Şekil 8'de görüldüğü üzere, robot hedef konum ile arasındaki mesafeyi sıfıra indirene kadar hareketine devam etmiştir. İkinci deneyde ise robot, aynı başlangıç pozisyonunda (0,0) başlatılmış, hedef nokta ise (-3,-2) olarak tanımlanmıştır. Robotun hedef konuma ilerlerken izlediği yol Şekil 9'da

gösterilmektedir.



Şekil 9: Deney 2 Robotun Yolu

Bu deneyde, robotun başlangıç pozisyonundaki yönü hedef noktaya ters olduğundan, önce hedefe doğru bir dönüş gerçekleştirmiş, ardından hedefe doğru hareket etmiştir. Robotun hedef nokta ile arasındaki mesafenin değişimi Şekil 10'da detaylı bir şekilde incelenebilir.



Şekil 10: Deney 2 Robotun hedef konuma uzaklığı

IV. SONUÇ

Bu çalışma, açık kaynak kodlu bir gezgin robot projesi geliştirilirken kullanılacak açık kaynak araçların incelenmesini amaçlamaktadır. Proje kapsamında öncelikle hem karasal hem de hava araçları üzerinde bir değerlendirme yapılmış, bu inceleme sonucunda karasal bir gezgin robotun kullanılmasına karar verilmiştir. Bu bağlamda, ROS tabanlı ve açık kaynak bir platform olan TurtleBot3 tercih edilmiştir.

Proje geliştirme sürecinin bir diğer aşamasında, uygun yazılım çerçevelerinin belirlenmesi gerekliliği ortaya çıkmıştır. Bu doğrultuda, açık kaynak kodlu olması ve geniş kullanıcı desteği nedeniyle ROS (Robot Operating System) tercih edilmiştir. Simülasyon ortamı için ise Webots ve Gazebo platformları değerlendirilmiş; uygulama gereksinimlerine uygun olarak Gazebo kullanılmıştır.

Çalışmada, robotun çevresel engelleri tespit etmesi ve bu engellerden kaçarak başlangıç konumundan hedef konuma ilerlemesi sağlanmıştır. Robotun bu işlevi gerçekleştirebilmesi için kamera görüntüleri üzerinden işlem

yapılmış ve görüntü işleme algoritmalarını çalıştırmak üzere açık kaynaklı bir kütüphane olan OpenCV kullanılmıştır. Robot, önündeki engelleri başarıyla tespit ederek, belirlenen hedef noktaya sorunsuz bir şekilde ulaşmıştır.

Projenin tüm aşamaları şekil ve grafiklerle detaylandırılmıştır. Bu çalışmada yalnızca RGB kamera kullanılmıştır; ancak ileriki çalışmalar kapsamında derinlik kameralarının da sisteme entegre edilmesi planlanmaktadır. Böylece, daha başarılı ve hassas sonuçlar elde edilmesi beklenmektedir. Ayrıca, kamera görüntülerinin işlenmesinde derin öğrenme ağlarının uygulanması düşünülmekte, bu sayede bilgisayarla görme performansının artırılması hedeflenmektedir.

KAYNAKLAR

- [1] S. Çaşka and A. Gayretli, "A survey of UAV/UGV collaborative systems," *CIE44 IMSS*, vol. 14, pp. 453–463, 2014.
- [2] Y. Lu, Z. Xue, G.-S. Xia, and L. Zhang, "A survey on vision-based UAV navigation," *Geo-spatial Inf. Sci.*, vol. 21, no. 1, pp. 21–32, 2018.
- [3] L. Fortunato and M. Galassi, "The case for free and open source software in research and scholarship," *Philos. Trans. R. Soc. A*, vol. 379, no. 2197, p. 20200079, 2021.
- [4] D. Riehle, "Free and Open Source Software," *Computer (Long Beach Calif.)*, vol. 57, no. 8, pp. 114–118, 2024.
- [5] A. Bekmez and K. Aram, "A Communication System for Dynamic Leader Selection in Distributed UAV Swarm Architecture," *Eur. J. Tech.*, vol. 14, no. 1, pp. 96–102.
- [6] A. Farley, J. Wang, and J. A. Marshall, "How to pick a mobile robot simulator: A quantitative comparison of CoppeliaSim, Gazebo, MORSE and Webots with a focus on accuracy of motion," *Simul. Model. Pract. Theory*, vol. 120, p. 102629, 2022.
- [7] F. M. López-Rodríguez and F. Cuesta, "Andruino-a1: Low-cost educational mobile robot based on android and arduino," *J. Intell. & Robot. Syst.*, vol. 81, no. 1, pp. 63–76, 2016.
- [8] D. Betancur-Vásquez, M. Mejia-Herrera, and J. S. Botero-Valencia, "Open source and open hardware mobile robot for developing applications in education and research," *HardwareX*, vol. 10, p. e00217, 2021.
- [9] S. Kohlbrecher, J. Meyer, T. Graber, K. Petersen, U. Klingauf, and O. Von Stryk, "Hector open source modules for autonomous mapping and navigation with rescue robots," in *RoboCup 2013: Robot World Cup XVII 17*, 2014, pp. 624–631.
- [10] K. Aram and A. Bekmez, "Leader-Follower Based Formation Control of Heterogeneous UAV-AGV Multi-Agent System," *Eur. J. Tech.*, vol. 13, no. 2, pp. 241–248, 2023.
- [11] A. Bekmez and K. Aram, "Three Dimensional Formation Control of Unmanned Aerial Vehicles in Obstacle Environments," *Balk. J. Electr. Comput. Eng.*, vol. 11, no. 4, pp. 387–394, 2023.
- [12] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," in *ICRA workshop on open source software*, 2009, p. 5.
- [13] L. Joseph and J. Cacace, *Mastering ROS for Robotics Programming: Design, build, and simulate complex robots using the Robot Operating System*. Packt Publishing Ltd, 2018.
- [14] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, "Robot operating system 2: Design, architecture, and uses in the wild," *Sci. Robot.*, vol. 7, no. 66, p. eabm6074, 2022.
- [15] O. Michel, "Cyberbotics Ltd. webots™: professional mobile robot simulation," *Int. J. Adv. Robot. Syst.*, vol. 1, no. 1, p. 5, 2004.
- [16] S. TIMOFTEI, B. Emilia, S. Anca, and S. Ovidiu, "Open-source software in robotics," *ACTA Tech. NAPOCENSIS-Series Appl. Math. Mech. Eng.*, vol. 61, no. 3, 2018.

Hyperloop Ekosistemi için Açık Kaynak Yazılım Mimarisi Modellenmesi

Open Source Software System Architecture Modelling for Hyperloop Ecosystem

Emrah BÖLÜK
Electronics Engineering
Gebze Technical University
Kocaeli, Türkiye
emrahblk00@gmail.com

Özetçe—Hyperloop teknolojisi yüksek hızlı kara taşımacılığı alanında devrim yaratmayı hedefleyen yeni nesil bir ulaşım yöntemi olarak karşımıza çıkmaktadır. Teknolojinin merkezinde vakumlu tüneller yer almaktadır. Tünellerin içerisindeki kapsüller ulaşımı sağlamaktadır. Yüksek hızlı demiryolu sistemlerinde sistem mimarisi ve kontrol konusu her zaman geliştirilebilir bir başlık olarak karşımıza çıkmaktadır. Hyperloop teknolojisinin mevcut potansiyelini koruması ve yeni bir ulaşım çağını başlatabilme ihtimali oldukça yüksektir. Bu nedenle sisteme yönelik açık kaynak yazılım modellerinin geliştirilmesi; bu teknolojinin daha ileri bir seviyeye gitmesi için oldukça kritik öneme sahiptir. Bu çalışmada model belirlenmiş bir hyperloop aracı için yazılım sistemi tasarımı gerçekleştirilmiştir. Önerilen yazılım modelinde ; güvenlik, enerji verimliliği, veri akışı, sensör entegrasyonu, otonom kontrol mekanizmaları ve iletişim altyapısı başlıkları ele alınmıştır. Ayrıca bu çalışma hyperloop teknolojisine genel bir bakış açısı sunmaktadır. Hyperloop teknolojisinde manyetik levitasyon ve tahrik kullanılır bu sayede çevre birimleri ile fiziksel temas ortadan kaldırılmıştır. Bu da kapsülün yüksek hız seviyelerine çıkmasını sağlamaktadır. Yüksek hızlı tren teknolojilerinde tepki süresi ve sistem karmaşıklığının yönetilebilir olması için yazılım mimarisi oldukça önemli bir rol oynamaktadır. Önerilen model “Araç Gömülü Yazılımı” ve “Kontrol Bilgisayarı Yazılımı” olarak ikiye ayrılmaktadır. Araç gömülü yazılım bölümünde; Çevresel izleme, Enerji yönetimi, Aktüatör (Eyleyici), Sensör, Navigasyon ve Haberleşme, Otonom kontrol birimleri bulunmaktadır. Kontrol bilgisayarı yazılım bölümünde; Sistem kontrolü akış, Arayüz, Haberleşme birimleri bulunmaktadır. Yazılım sistemi modeli tamamlandıktan sonra hyperloop teknolojisine ve gelecekte yapılabilecek sistem geliştirmelerine değinilerek çalışma tamamlanmıştır.

Anahtar Kelimeler: Hyperloop, Kapsül, Manyetik Levitasyon, Gömülü Yazılım

Abstract-- Hyperloop technology is a new generation of transportation that aims to revolutionize high-speed ground travel. At the heart of the technology are vacuum tunnels. The capsules inside the tunnels provide the transport. In high-speed rail systems, system architecture and control are always an evolving issue. Hyperloop technology is likely to maintain its current potential and usher in a new transport era. Therefore, developing open-source software models for the system is crucial for this technology to move to a more advanced level. In this study, a software system design for a modeled Hyperloop vehicle has been carried out. The proposed software model discusses safety, energy efficiency, data flow, sensor integration, autonomous control mechanisms, and communication infrastructure. In addition, this study provides an overview of

hyperloop technology. Hyperloop technology uses magnetic levitation and propulsion, eliminating physical contact with the peripherals. This allows the capsule to reach high speeds. Software architecture plays an important role in high-speed train technologies to manage response time and system complexity. The proposed model is divided into “Vehicle Embedded Software” and “Control Computer Software”. In the vehicle embedded software section; environmental monitoring, energy management, actuator, sensor, navigation and communication, and autonomous control units are included. In the control computer software section; system control flow, interface, and communication units are available. After completing the software system model, the study concludes by mentioning the Hyperloop technology and system developments that can be made in the future.

Keywords: Hyperloop, Pod, Magnetic Levitation, Embedded Software

I. GİRİŞ

A. Hyperloop Genel Bakış

Ulaşım konusu geçmişten günümüze kadar medeniyetin gelişiminde her dönemin önemli yapı taşlarından biri olmuştur. Bu bağlamda iletişim teknolojileri de insan hayatını etkileyen en önemli faktörlerin başında gelmektedir. Bu iki alanın ortak noktasında geliştirilen teknolojiler tarihte her zaman kritik bir değere sahip olmuştur[1]. Hyperloop teknolojisinin temel konsepti günümüz teknolojisiyle donatılmış olsa da çalışmaların başlangıç noktası 20. Yy’ın başlarına dayanmaktadır. 1910 yılında Amerikalı mühendis Robert Goddard tarafından vakumlu tren fikri ortaya atılmıştır fakat o dönem ki teknolojik altyapı yetersizliği nedeniyle hayata geçirilememiştir. Daha sonrasında 1970’lerde Rand Corporation Firmasında yapılan düşük basınçlı tüplerle çalışan hızlı tren çalışmaları bu teknolojideki önemli hususlardan biri olmuştur. 2013 yılında SpaceX CEO’su Elon Musk tarafından Hyperloop alpha adlı bir teknik rapor yayınlamıştır. Bu çalışma Los Angeles ve San Francisco arası yolcu taşıma potansiyeline sahip bir pod tasarımını göre geliştirilmiştir. [2] Bu kapsül tasarımı manyetik levitasyon ve elektrikli tiki sistemine sahip ilk çalışma örneğidir. Bu rapor Hyperloop sisteminin maliyet ve uygulanabilirlik başta olmak üzere tüm teknik açılardan yapılabileceğini göstererek çalışmaların hızlanmasına vesile olmuştur. Bu çalışmada da hyperloop teknolojisinin sistem haberleşmesine dair bir model tasarımı önerilecektir.

B. Model Tanımı

Hyperloop sistemleri; otonom kontrol, güvenlik mekanizması, sensör veri entegrasyonu, enerji yönetimi ve ana bilgisayar(sunucu) iletişimi gibi gereksinimleri karşılayabilecek bir yazılım mimarisine ihtiyaç duyar. İnsan taşımacılığında kullanılması planlanan hyperloop sistemleri oldukça yüksek standartlara ve düzene tabiidir.[3] Hyperloop'un uluslararası ulaşım sektöründe potansiyeline ulaşması için fiziksel altyapıyla birlikte güçlü bir yazılım ekosistemine ihtiyaç duyulmaktadır. Bu çalışmada yazılım mimarisi modellerine ilişkin tüm teknik detaylar ve kavramsal temeller atılarak bu yeni taşımacılık sistemine yönelik gelecek araştırmalara yardımcı olmak hedeflenmektedir. Yazılım sistemi aşağıdaki şekilde görüleceği gibi 2 ana alt başlıkta incelenmektedir.



Fig. 1. Yazılım Mimarisi Blok Genel Bakış

II. YAZILIM MİMARİSİ

A. Kontrol Bilgisayarı

Hyperloop sisteminde gerçek zamanlı veri akışı ve ana bilgisayar haberleşmesi için kontrol arayüz sistemi oldukça kritik bir yazılım katmanıdır. Kapsülün uzaktan erişime açık ve denetlenebilir olması kullanıcı güvenliği için esas faktördür. Hyperloop kontrol sistemi, temelinde ölçeklenebilmesi ve sistemin soyutlaştırılmasının kolaylaşması için Master-Slave tasarımı kullanılacaktır. Bu sistemin güvenli kontrol altyapısı ve yüksek performans vermesi hedeflenmektedir. Sistemin ana sunucusu Kontrol bilgisayarıdır. Genel sistem işleyişinden sorumludur. Kontrol bilgisayarı ve hyperloop kapsülü içerisinde kullanılan master bilgisayar QUIC protokolü kullanılarak wifi ile haberleşir. UDP tabanlı bir yapı olan QUIC protokolü düşük gecikmeli çalışmaktadır [4], ayrıca kuyruk mimarisine sahip olduğu için sıralı ve hızlı bir veri aktarım imkanı sunmaktadır. Kontrol bilgisayarı kapsülün genel işleyişini ve durumunu denetler, gerektiği durumda manuel müdahale imkanı sağlar. Ayrıca kapsül içerisinde bulunan tüm çevresel birimleri denetleyerek alt sistemlerle ve işlem basamaklarıyla olan bağlantıları sağlamaktadır. [5]

Arayüz bloğu sistem operatörlerinin kapsülü takip edebilmesi ve parametrelerin okunması için tasarlanan web tabanlı bir yapıdır. Arayüz bloğunda aşağıdaki temel yapılar bulunacak, işlevler yerine getirilecektir.



Fig. 2. Tasarlanan Örnek Arayüz

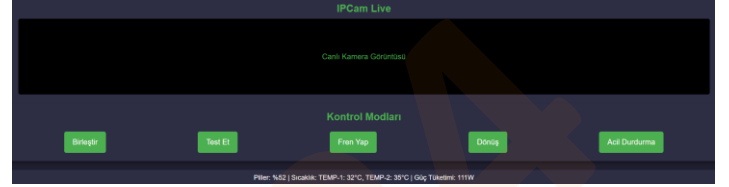


Fig. 3. Tasarlanan Örnek Arayüz

3D 2D Görselleştirme: Tünel içerisinde bulunan kapsülün çevresel durumu, rotası hedefe olan uzaklığı gibi bilgiler 3D harita üzerinde kullanıcıya sunulabilir, ayrıca daha az karmaşık olması adına 2D görselleştirme yapılabilir.

Manuel ve Otonom Müdahale: Kapsül otonom hareket algoritmalarını uygulayarak hareket edebilir olası aksi bir durumda arayüz üzerinden manuel müdahale edilerek tehlikeli durumların önüne geçilir. Tüm bunlar panelden takip edilebilir.

Gerçek Zamanlı Veri Paneli: kapsülün, konum, sıcaklık, basınç, hız gibi çevresel sensörlerden gelen verileri arayüz üzerinden görünür halde olacaktır. Bu sayede operatörün sistem dinamiklerine ulaşımı oldukça erişilebilir olacaktır.

Veri Kaydı: Kapsül hareket başladığı andan itibaren tüm veriler kayıt altına alınarak gelişebilecek olayların kronolojik olarak gözlemlenmesi- takip edilmesi sağlanacaktır. Bu sayede hata tespiti ve analizi yapılarak sistem iyileştirmeleri gerçekleştirilecektir. Veriler bulut ve yerel depolamada belli periyotlarda yedeklenecektir böylelikle hem kontrol bilgisayarıdan hem de uzaktan erişim imkanı sunulacaktır.

Şifreleme: Sistemin veri güvenliği performans açısından oldukça kritik öneme sahiptir. QUIC protokolü dahilinde bulunan TLS sistemi veri güvenliğini sağlamaktadır. [6] Veri iletim sisteminde kontrol bilgisayarı – master bilgisayarı – slave cihazları arasında (end to end) uçtan uca şifreleme kullanılacaktır. Veri kayıt dosyalarında AES algoritması, Şifreli verilerin doğrulanmasında hash algoritmaları ve sisteme giriş yapan operatörlerde çift aşamalı kimlik doğrulanması uygulanacaktır.

B. Araç Gömülü Yazılımı

Araç gömülü yazılımı sistemin en kritik bileşenlerinden biridir. [7] Bu sistem kapsülün kontrolü, çevresel algılama, güvenlik yönetimi gibi sistemlerdeki işlevleri yerine getirir. Gömülü yazılım, araç içerisinde bulunan elektronik ve mekanik bileşenlerin kontrol edilmesi için düşük gecikme süreli kesintisiz komut iletimi sağlayarak sistemin sorunsuz şekilde ilerlemesini sağlar.

Tasarlanan sistemde kontrol bilgisayarı ve gömülü yazılım bilgisayarı arasında Wifi üzerinden UDP tabanlı QUIC protokolü kullanılır. Gömülü yazılım bilgisayarı araç içerisindeki tüm çevresel birimlere kontrol komutlarını

gönderir aynı zamanda Ana sunucu bilgisayarından gelen manuel komutları iletir. Çevresel birimler ve gömülü yazılım bilgisayarı arasından çift yönlü iletişim vardır. Eski veriler ve güncel komutlar düzenli olarak gönderilir-alınır. Bu sayede otonom harekete devam edilirken, bilgiler ana sunucuya aktarılarak veri akışı sağlanmış olur. 6 alt birimden oluşmaktadır.

1. Çevresel İzleme Birimleri:

Kapsülün tünel içerisindeki durumu ve çevresel faktörleri bu birimdeki sensörler vasıtasıyla takip edilir. Kontrol bilgisayarına hava kalitesi partikül yoğunluğu nem seviyesi gibi veriler iletir. Olumsuz hava koşulları olması durumunda gerekli kontrol ve hareket mekanizmaları devreye girer.

Tüpün hareketi boyunca sıcaklık basınç, hız ve tüm diğer değişkenler takip edilir. Aksi bir durumda acil durum otonom sürüş algoritmalarından gerekli olanlar devreye girer.

2. Enerji Yönetimi Birimi:

Sistem enerji paketi ana batarya, acil durum bataryası ve elektronik sistem bataryasından oluşmaktadır.

BMS Sistemi: Kapsül enerji tüketimi, yönetimi ve verimliliğini denetlemek, iyileştirmek için enerji yönetim birini oluşturulmuştur. Bataryanın voltajı, sıcaklığı ve şarj durumu takip edilmektedir. Enerji akışını yönlendirmek ve sınırlamak için gelen veriler kontrol edilerek BMS'ye uygun komutlar verilir.

3. Aktüatör (Eyleyici) Birimleri:

Kapsülün tünel içerisindeki hareketini yerine getirmekle görevli çevresel elemanlardan oluşmaktadır.

Manyetik Levitasyon ve İtki Sistemi: Aracın tünel içerisindeki ray ile temasını kesen ve ileri yönlü hareketini sağlayan sistemdir. Manyetik disklerden ve lineer motordan oluşmaktadır. Otonom sürüş algoritmasından gelen komutlara göre kapsülün hareketini sağlar.

Frenleme Sistemi: Kapsülün fren yapabilmesini sağlamak için basınçlı gaz, mekanik veya elektrik temelli fren sistemleri oluşturulabilir. Bu sayede aracın hızının düşürülmesi veya yolculuk sonunda durması sağlanmaktadır. Frenleme komutlarıyla beraber güvenlik prosedürleri de devreye girer, bu komutlar otonom sürüş algoritması veya sunucu tarafından manuel olarak verilebilir.

4. Sensör Birimleri:

Kapsül içerisinde ve çevresinde, durum analizi ve sistem bilgilerinin toplanması için sensörler bulunmaktadır. Sensörlerden gelen veriler gerçek zamanlı olarak işlenerek, otonom sürüş algoritmasının kararlarına yardımcı olur aynı zamanda tüm veriler arayüz üzerinden sunucuya aktarılır. Kullanılan başlıca sensörler şu şekildedir:

- 360 derece eğim sensörü
- İvmeölçer
- Hız sensörü
- Akü, tekerlek, iç ortam basınç sensörleri
- Fotoelektrik mesafe sensörleri
- Motor, tekerlek, elektronik donanım, sıcaklık sensörü
- Akım sensörü
- Lidar ve kamera

5. Navigasyon ve Haberleşme Birimi:

Kapsülün hareketi ve rota takibini yapabilmesi ayrıca diğer sistemlerin birbiriyle haberleşmesini sağlamaktadır. GNNS ve lazer işaretleyiciler modüller kullanılarak kapsülün pozisyonu tayin edilecektir. Kontrol bilgisayarındaki otonom sürüş algoritmaları verileri optimize ederek navigasyonu tamamlar. Kapsülde farklı iletişim paketleri kullanılarak hem cihazlar arası hem de merkezi kontrol ile iletişim sağlanır.

CAN-Bus Protokolü:Gömülü yazılım bilgisayarı ve alt sistemler arası veri akışını sağlamak amacıyla kullanılır. Sensörlerden gelen veriler kontrol bilgisayarına aktarılır ardından tekrar komutlar bu ağ aracılığıyla iletir.

QUIC Protokolü: Ana sunucu bilgisayarı ile gömülü yazılım bilgisayarı arasındaki iletişimde kullanılır. Bu bağlantı wifi üzerinden sağlanır. Aynı bağlantı üzerinde paralel işleme yaparak birden fazla veri akışının iletilmesi ve işlenmesine olanak sağlar. Uçtan uca şifreleme yaparak verilerin güvenli şekilde aktarılmasını sağlar aynı zamanda düşük gecikme ve hızlı bağlantı sağlaması bu noktada önemli katkı sağlamaktadır.

Veri Paketi Yapısı: Sisteme iletilen veri paketleri türü, kimliği ve zaman damgasına göre ayrılmaktadır ardından veriler veya komutlar elde edilmektedir.

Kuyruk Mimarisi: Bu yapı veri akışını sistemin gereksinim sıralamasına göre yöneterek işlemlerin gerçekleştirilmesine olanak tanır. Kuyruk yapısı düşük orta ve yüksek derece olmak üzere 3'e ayrılır. Yüksek öncelikli kuyruk acil durum senaryolarında geçerli olmaktadır. Orta öncelikli kuyruk aracın hareketi ve sensör verileri gibi rutin senaryoları kapsamaktadır. Düşük öncelikli kuyrukta ise aracın hareketinden bağımsız olan sensör verilerini ve performans iyileştirmelerini kapsamaktadır.

6. Otonom Kontrol Birimi:

Kapsülün kararlı, enerji verimliliği yüksek ve güvenli bir şekilde hareket etmesini sağlamak için otonom kontrol algoritmaları oldukça önemlidir. Bu nedenle çeşitli algoritmalar kullanılarak gerçek zamanlı komutlar gerçekleştirilir.[8]



Fig. 4. Otonom Kontrol Birimi Alt Dalları

Belirli Sonlu Otomat (DFA) : Kapsülün durumlar arası geçişini sağlar. Başlatma, hızlanma, yavaşlama, durma, acil durum gibi kritik komutlar arası geçişi sağlar. Operasyonel durumları yönetir.

Model Öngörülü Kontrol (MPC): Gerçek zamanlı verilere göre kapsülün hareketlerini inceler, gelecek hareketlerine dair tahminler oluşturur sistemi optimize eder.

Oransal-İntegral-Türevsel (PID) Kontrol: Sensörlerden gelen hız, konum gibi dinamik parametreleri kontrol eder. Gelen verilere göre işleme yaparak gerekli düzeltmeleri yapar.

Pekiştirmeli Öğrenme (RL) Algoritmaları: Kapsülün geçmiş verilerine dayanarak yeni bir model öngörüsü sunar, farklı test senaryolarını öğrenerek en optimalini sunmaya

çalışır. Sistem dinamiklerine kolay optimizasyon sağlayarak karmaşık duumlarda kontrol stratejisi geliştirebilir.

Kayan Kipli Kontrol (SMC) : Ekosistem değişkenlerinde meydana gelen bozulma ve değişimlere yönelik kontrol algoritması oluşturur. Ani değişimlere karşı hızlı reaksiyon sağlar.

Yapay Sinir Ağları : Kapsül içerisinde ve çevresinde gelen büyük boyutlu veriyi optimal hızda analiz ederek kontrol komutlarının verilmesine yardımcı olur. Öğrenme kapasitesi yüksektir, model değişikliklerine karşı hızlı oldukça iyidir.

Yukarıda önerilen otonom kontrol algoritmaları, araç hareketinde oldukça önemli olan konum, hız, güvenlik, enerji gibi parametrelerin düzenlenmesi ve değerlendirilmesinde rol alarak otonom hareket işleyişini sağlar, ayrıca sistem iyileştirmelerinde de kritik rol alır.

III. SONUÇ VE GELECEK ÇALIŞMALAR

Bu çalışmada Hyperloop teknolojisine dair gelişmiş bir yazılım mimarisi modellemesi yapılmıştır. Çalışma esnasında araç gömülü yazılımı ve kontrol bilgisayarı yazılımı ana alt başlıkları oluşturmuşlardır. Bu başlıklar çerçevesinde altyapı ve ulaşım kriterleri göz önünde bulundurularak model geliştirilmesi tamamlanmıştır. Kullanılan modelde ; güvenlik, enerji verimliliği, veri akışı, sensör entegrasyonu, otonom kontrol mekanizmalarının yüksek hızda düşük gecikmeyle çalışması hedeflenmiştir. Gömülü yazılım sisteminde, çevresel birimlerden gelen gerçek zamanlı veriler analiz edilerek otonom kontrol algoritmalarının daha doğru komutlar vermesi sağlanmıştır. Kontrol bilgisayarı yazılımı kısmında ise kapsül içerisinde bulunan sensörlerden alınan tüm veriler gömülü yazılım bilgisayarı tarafından ana sonucu bilgisayara iletilmektedir. Bu sayede araç durumunun görselleştirilmesi ve verilerin kayıt altına alınması sağlanmıştır. Ayrıca olası acil durumlarda ana bilgisayar üzerinden müdahale imkanı tanınmasına olanak sunulmuştur.

Bununla birlikte, hyperloop teknolojisindeki geliştirmeler ve yapılan çalışmalar her geçen artmaktadır. Teknolojinin gelişmesiyle birlikte farklı yazılım sistemleri geliştirilmeye ihtiyaç duyulacaktır. Bu noktada olası gelecek çalışma konuları şu şekildedir:

Sensör Kullanımı: LIDAR, manyetik alan ölçüm sensörü, stabilizasyon-titreşim sensörü, termal kameralar gibi farklı teknolojilerin entegrasyonu, navigasyon ve güvenlik başta olmak üzere sistemde oldukça faydalı olabilir.

İletişim Protokollerinin Değiştirilmesi: Şuan mevcut protokoller çerçevesinde QUIC protokolü tercih edilmiştir fakat ileride 6G ve üzeri teknolojilerin yaygınlaşmasıyla daha hızlı ve düşük gecikme süreli bir protokol kullanılabilir veya mevcut protokolle bu yönde iyileştirmeler sağlanabilir. Bu sayede sistemin daha hızlı ve güvenli hale gelmesine fayda sağlanabilir.

Donanım Arıza Durumu ve Yedekliliği: Sistemde kullanılan elektronik bileşenlerin arıza durumu- toleransı ve yedekliliği konusu oldukça geliştirilmeye açıktır. Bu noktada gerçek dünya testleri ve donanım döngü testleri (HIL) artırılarak sistem iyileştirmeleri yapılabilir.

Gelişmiş Algoritmaların Sistem Entegrasyonu: Makine öğrenmesi ve yapay zeka tabanlı geliştirilmiş modeller kullanılarak, otonom kontrol algoritmaların daha hassas ve doğru tepkiler vermesi sağlanabilir. Derin öğrenme modeli yeterli veri ile eğitildikten sonra araç acil durumlarında da otonom kontrol mekanizmasının karar yetisi geliştirilebilir.

Yukarıdaki konular başta olmak üzere mevcut teknolojide bir çok noktada yenilik sağlanarak sistem geliştirilmesi yapılması oldukça mümkündür. Bu çalışma gelecekte yapılacak çalışmalara bir referans, yol haritası olma niteliği taşımaktadır.

KAYNAKLAR

- [1] M. Liem and V. B. Mendiratta, "Mission critical communication networks for railways," Bell. Labs Techn. J., vol. 16, no. 3, pp. 29–46, 2011
- [2] E. Musk, Hyperloop Alpha, SpaceX, Hawthorne, CA, USA, 2013.
- [3] Schafer, A. and Victor, D. G., "The future mobility of the world population," Transportation Research Part A: Policy and Practice, Vol. 34, No. 3, 2000, pp. 171–205. Cited on p. 15
- [4] Google. (t.y.). QUIC: UDP üzerinde çoklu akış taşıma protokolü. Erişim adresi: <https://www.chromium.org/quic> (Erişim tarihi: 15 Kasım 2024).
- [5] Hyperloop Transportation Technology. (t.y.). Hyperloop Transportation Technology. Erişim adresi: <https://www.hyperlooptt.com/> (Erişim tarihi: 15 Kasım 2024).
- [6] H. Yu, J. Li, and J. Wu, "Integrating TLS 1.3 into QUIC for secure and reliable data transmission," IEEE Transactions on Networking, 2022.
- [7] K. van Goeverden, D. Milakis, M. Janic, and R. Konings, "Analysis and modelling of performances of the HL (Hyperloop) transport system," Eur. Transp. Res. Rev., vol. 10, no. 2, p. 41, 2018.
- [8] S. Jeker, "Hyperloop network design," M.S. thesis, Comput. Eng. Netw. Lab., ETH Zürich, Zürich, Switzerland, 2019.

Modernizing Healthcare Systems: An Open-Source Provider Portal with AI and FHIR Integration

Cihan Topuz
Computer Science
The Georgia Institute of Technology
Atlanta, USA
ctopuz3@gatech.edu

Fahim Bin-Afzal Al-Ashrafi
Computer Science
The Georgia Institute of Technology
Atlanta, USA
fahim.rafi@gatech.edu

Benedicto V Da Silva
Computer Science
The Georgia Institute of Technology
Atlanta, USA
ben.dasilva@gatech.edu

Abstract—This paper presents an open-source healthcare portal designed to address the challenges of health data interoperability and clinical decision-making in modern medicine. As healthcare systems become more reliant on complex technology, integrating disparate health data has never been more important. The same capability is greatly enhanced by the FHIR standards integration with artificial intelligence (AI) coordinating to act through the portal. AI embedded within the portal automatically synthesizes patient encounters, medication histories, clinical notes, and procedural records into brief summaries designed to deliver actionable insights directly to providers. This may be more than just another integrated technological innovation or even a simple tool, but rather a transformative shift creating a more integrated and user-friendly healthcare system. This paper demonstrates the potential for open-source technologies to fundamentally change the way providers, patients, and others can interact with health information and each other. These future possibilities for patient care are refined further through open source, and when enacted, will lead to improved clinical outcomes and experience.

Keywords—Open-source healthcare, FHIR, AI in healthcare, data interoperability, provider portal

I. INTRODUCTION

A. Background

As healthcare advances at an unprecedented rate, the influx of patient data has rapidly grown, forcing providers to interpret and act on significant amounts of information in a limited amount of time. With the arrival of EHR systems and FHIR standards, we have seen a lot of things improve, and both structured and unstructured data are available to assist providers. Yet, these systems have considerable ergonomic shortcomings, resulting in inefficiency and subsuming critical data under layers of complexity. This is an aggravation when trying to access and sift through information in a quick and timely manner.

Providers must depend on sound systems to assess this data properly. However, the majority of healthcare organizations do not possess the resources or expertise to create tools that utilize advanced technology such as Artificial Intelligence (AI) to provide a holistic view of FHIR data. This leads to providers struggling with disparate workflows and a hindrance to clinical decision-making with added administrative challenges. According to Junttila, physicians emphasize the importance of a single service that provides

access to the information needed to make well-informed decisions and save time. [1]

This paper addresses these challenges by combining FHIR standards with AI-driven insights to transform raw healthcare data into actionable intelligence. By bridging the gap between data accessibility and usability, the proposed portal aims to alleviate administrative burdens and enhance the quality and efficiency of patient care. Specifically, the portal's objectives include:

- Abstracting patients' medical histories and encounters into concise summaries.
- Identifying meaningful trends in clinical notes, medications, and procedures.
- Accelerating decision-making by providing clear, actionable insights.

Through these capabilities, this paper aims to empower healthcare providers and modernize the way they interact with patient data and allowing providers to spend more time with the patient.

B. Problem Statement

Modern healthcare systems face a number of challenges, evolving from improper data communication and inefficient decision-making processes. One of the core questions that arises here is how to create an information infrastructure that is both standardized and flexible.[2] The rush towards digitalization and globalization of healthcare systems calls for coherent and synergistic efforts towards data governance, or otherwise prevailing tensions within the sectors will compound. This increase will make the coordination between the different disciplines involved in providing care even more difficult.

Another cause for concern is constant complaints from front-line physicians about the existing information systems. Studies highlight how fragmented systems force providers to spend valuable time navigating multiple platforms to access essential information [1]. Also, according to the Medscape Physician Compensation Report 2023, physicians spend an average of 15.5 hours per week on paperwork and administrative duties. [6] This disintegration creates a great barrier to accessing key information that the healthcare team considers necessary to ensure high-quality patient care.

A proposed solution to this could be the creation of an open-source provider portal that is built and improved upon with mechanisms that allow advanced technologies like AI and FHIR to be integrated into them. Given that data are managed at different IT stores and locations, implementation of such a portal should also facilitate the smooth exchange of data, increasing efficiency and ultimately improving outcomes. This initiative has the potential to help overcome the barriers identified by promoting the sharing and use of healthcare data and supporting healthcare providers' decision-making. This unified platform between AI and FHIR is a leap toward a more consolidated and agile healthcare delivery system, benefiting practitioners and patients alike.

C. Objectives

In the current environment of healthcare delivery system modernization, there is an even sharper focus on the need for clarity around the objectives of each aspect of modernization to provide benefits by ensuring efficiency and effectiveness. One of the central goals in today's healthcare environment is to improve decision-making in an environment where healthcare providers continue to be frustrated with current information systems. These systems typically do not sufficiently present the needs of providers or accommodate clinical workflows' complexities. [1] Physicians face immense issues due to the fragmentation of data among many disparate silos. Such fragmentation adds a devastating challenge to timely and informed decision-making and, ultimately, hinders the quality of patient care.

In light of these challenges, the purpose of this paper is to establish a seamless flow of information among the diverse stakeholders within the healthcare ecosystem. This will be made possible by establishing a provider portal that leverages open-source technology with AI and the Fast Healthcare Interoperability Resources (FHIR) standard. The incorporation of FHIR-compliant data presents a unique opportunity to leverage machine learning algorithms, thereby enhancing analytical capabilities and facilitating the extraction of meaningful insights. Healthcare organizations can leverage robust data infrastructures, such as the Health Information Gateway and Exchange (HINGE) framework, to strategically collect and structure distribution datasets such as clinical, genomic, and imaging data and integrate them into collaborative knowledge bases. Integrating these knowledge bases could prove valuable in the efforts to advance personalized care that ultimately improves care outcomes for patients.[4]

In addition, BETTER (the Horizon Europe project) [5] shows the potential of applying machine learning approaches to FHIR data, and how such applications can help ensure interoperability and improve the quality of clinical decision-making. On the other hand, Telehealth solutions, augmented with AI and adhering to FHIR standards, have demonstrated notable success in managing chronic diseases, as evidenced by Chen. [3] This underscores the fact that telehealth services are not merely a theoretical concept but are increasingly becoming a tangible reality with the potential to fundamentally transform healthcare delivery. While this project focuses on provider-oriented insights, it represents an

initial step toward broader applications, such as connecting community healthcare, home care, and hospitals.

Altogether, this initiative aims to enhance transparency, usability, and trust in health information systems within the healthcare space, which ultimately supports the dynamic evolution of data-driven medicine. Of course, interfaces like Quasar also matter a lot as businesses build on top of Spring Boot and other open-source technologies, which generate the demand for cost-efficient healthcare tech evolution.

D. Literature Review

As technology is advancing rapidly, healthcare systems have also seen a paradigm shift towards the usage of open-source solutions. Multiple studies highlight the promising effects of these platforms to improve operations and care of patients. According to Healthcare Innovation, open-source solutions can lower implementation costs drastically and at the same time promote collaboration among all involved, which is necessary to improve global healthcare systems. [9]

On the other hand, the role of AI is also significant in improving patient care and operational efficiency in the race of digital transformation in healthcare systems. Using data mining, health trend prediction, and resource optimization, AI helps provide personalized treatment for patients. AI insights, fused with the power of FHIR standardization, create an enormous potential to advance clinical outcomes and operational efficiencies. [8] As such, the literature consistently underscores the potential of these technologies to modernize healthcare delivery, ultimately improving outcomes and patient satisfaction across various healthcare settings.

Further, the literature shows the radical potential of open-source platforms, AI, and interoperability standards to either update or revolutionize healthcare systems. They improve the accessibility and consistency of data, together overcoming grave inefficiencies in the delivery of healthcare. Their ability to improve clinical outcomes, operational efficiency, and patient satisfaction through continued investigation brings the full potential of these innovations to diverse healthcare settings.

II. METHODOLOGY

The methodology of this paper provides for a scalable, efficient, and secure healthcare portal. This enables real-time interoperability and actionable insights for healthcare providers utilizing modular architecture, synthetic data, and advanced artificial intelligence technologies.

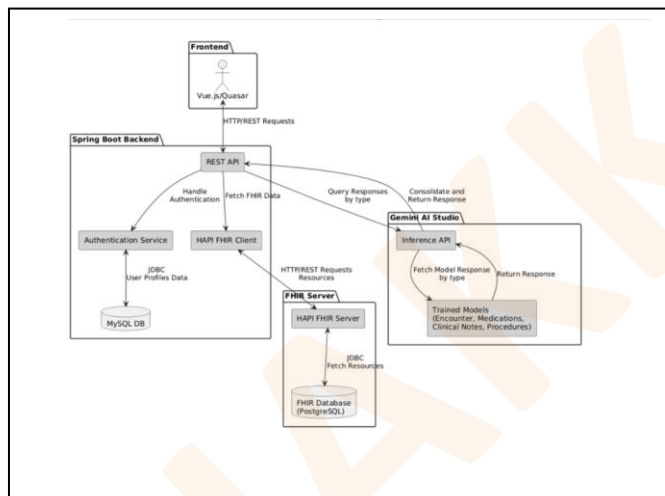
A. System Architecture

The architecture of this solution represents the backbone of the project, seamlessly integrating multiple components to deliver a smooth and reliable experience for healthcare providers. Furthermore, the system is capable of integrating with a variety of software solutions, including custom applications and third-party systems developed by alternative health solution providers, thereby enhancing the real-time processing of healthcare information. This architecture is

poised to ensure the efficiency of processes, the maintainability of code, the security of data, and the utilization of an open-source stack, ultimately delivering actionable insights to healthcare professionals. The Basic Design Principles of Architecture can be examined in three steps as follows.

- **Modularity:** Every part of the system, such as frontend, backend, FHIR server, and AI models, functions in its own way yet coherently. This modular approach eases maintenance and allows for future expansion.
- **Scalability:** The architecture supports the integration of new AI models or functions without disturbing existing systems. It is adaptable to changing healthcare needs.
- **Security:** Solid authentication mechanisms and secure database interactions protect sensitive patient information, maintaining compliance with healthcare regulations.

The system follows a modular design at a high level, ensuring scalability, maintainability, and robust functionality.



a) Front-end Architecture:

The front-end component, developed using Vue.js and Quasar, makes it visually comprehensible for healthcare service providers. It provides an interactive and user-centric platform, which minimizes attempts to contact the system on behalf of others. Communication with the backend is achieved using secure REST APIs to guarantee the timely exchange of data and message delivery reliability. For example, healthcare providers can rapidly learn from their patient's medication history. Assuming a patient calls in one day needing his medication records, the front-end:

- Communicates with the backend to collect the required information.
- Sends an HTTP request stating which sort of data is being looked for.

Through this arrangement, users and systems interact seamlessly with effects that are structured effectively according to data. When they pick out something, such as a piece of medication history, the system does all the work for them quickly and accurately.

b) Back-end Architecture:

The backend uses Spring Boot as a managing body that combines front-end, data resources, and artificial intelligence (AI) systems. It undertakes tasks essential to providing interactive applications like data integration, classification-based switching, and display interactions with AI models, such as reconfiguring their settings through iterative learning experience mode. The main components of the backend include:

- According to REST API principles, it receives requests from the front end, performs necessary operations, and returns the responses for the following functions.
- Automatically creates a schema based on metadata received during start-up.
- Draws or defines HTTP routes (keeping state between requests).

In dealing with requests of this type, the HAPI FHIR client talks to the FHIR server, keeping healthcare standards. Patient data is collected from various places in the FHIR server.' By its architecture, the authentication service provides secure access according to certain rules, while session data and user profiles are retained by a PostgreSQL database. Type-based routing permits the backend to use its dynamic knowledge about the nature of requested data (e.g., clinical notes, procedures) and meta-information integrated to aid inference on clinical narratives, sending it directly to the AI model for processing the content.

If the front end requires a user's clinical note summary data, then:

- The back end accepts the request, fetches the corresponding patient's data from the FHIR Server, and sends this information to the appropriate AI Model in Gemini AI Studio.

- The backend forwards the result to the front for real-time imaging.

c) FHIR Server (HAPI FHIR) :

HAPI FHIR is a complete implementation of the HL7 FHIR standard for healthcare interoperability in Java. The FHIR server is the primary repository for patient data, including encounters, medications, clinical notes, and procedures. It ensures interoperability by adhering to FHIR standards, enabling seamless data integration and retrieval. [10]

d) AI Integration (Gemini AI Studio) :

The intelligence layer (on the system) is governed by Gemini AI Studio, which is responsible for training AI models suited for particular healthcare uses. These AI models are capable of generating specialized insights for medical providers, and with the backend, the AI Studio is interfaced by an inference API supporting type-based routing of requests to appropriate models.

Particular AI Models:

- Encounters Model
- Medications Model
- Procedures Model
- Clinical Notes Model

Process Example:

- A request for a medication summary is sent to the Medications model.
- AI model processes data and generates a summary of the information.
- The backend collates the AI-generated output and transmits it to the front end for visualization.

Therefore, by implementing this architectural framework, healthcare practitioners' processes of meeting with patient data can be made secure, enhancing operations. On another level, the design has been chosen to prepare healthcare for the future: the design is contained within modular and open-source technology.

The portal is driven by synthetic healthcare data using Synthea, an open-source synthetic patient generator that produces semantically organized patient data for the purpose of FHIR-compliant data generation. This data mimics actual patient data from the real world (encounters, medications, clinical notes, and procedures). By hosting the store on a HAPI FHIR server, all providers can adhere to interoperability standards, and seamless integration into the back end.

B. Data Sources

In this project, the training of artificial intelligence (AI) is conducted in a systematic manner, and the essential elements include:

- Synthetic Data Generation: Patient data was generated with a focus on cardiovascular health incidents.
- Training Dataset Creation: A total of 107 input-output pairs to represent patients in a concise form were created from the output data records.
- Fine-tuning the AI Models: The preparations mentioned above were foundational material for calibrating the four specialized AI models available within Google Gemini Studio. Each model was purposefully designed to handle a certain form of data: the Patient Encounters, Medications, Procedures, and Clinical Notes model.

C. Data Processing

The pipeline is built to process the data and route requests appropriately to get the insights that matter. The backend is built with type-based routing, where we can intelligently know the type of data being requested from the core (i.e., whether it is encounters, medications, etc.) and then route it to the relevant AI models.

AI Models: The four specialized Gemini models were fine-tuned for different sets of data:

- Patient Encounters
- Medications
- Procedures
- Clinical Notes

After the request is made from the frontend, it is sent to the backend, where instances of data can be classified. For example, a query regarding "medications" will cause the backend to forward the request to the Medications AI model. The model takes this input and runs its magic, giving a summary of the text. Then, the processed data is securely returned to the frontend for visualization in a near real-time manner. On the other hand, The research explored several models, such as Claude 3.5 Sonnet and Gemini 1.5 Flash. Sonnet was initially used to develop accurate summary outputs, and these outputs formed the basis for fine-tuning Gemini models, which were selected for their precision and efficiency. This portal incorporates advanced AI models with the artificial creation of language operations, it can also be used at scale, providing healthcare professionals with precise, personalized, and immediately actionable details. This empowers them to make the best healthcare decisions for themselves and their patients.

III. SYSTEM DESIGN

The architectural framework supporting this system exhibits a robust, modular, and extremely scalable design. This framework effectively deals with the crucial healthcare data interoperability issues and clinical decision-making processes. It is both flexible and cost-effective, utilizing state of the art open-source technologies to serve the immediate

needs of medical providers while at the same time laying a solid foundation for future developments in this field.

A. Front-end Architecture

The front-end interface of our platform is built on the Quasar framework, which is constantly being updated and improved. [13] It has been designed to give healthcare institutions a highly interactive interface that is easy to use and fast to respond. Key features of the front-end architecture include:

- **Vue.js:** A JavaScript framework for creating user interfaces.
- **Quasar:** Vue.js-based framework for developing interactive web applications.
- **Apache ECharts:** A JavaScript charting library for data visualization.
- **OIDC Client:** A JavaScript library for OpenID Connect and OAuth2 authentication.
- **Pinia:** Vue.js store library to manage the state of the application.
- **Patient Management:** The system involves a deep, simple patient list directly queried from the FHIR (Fast Healthcare Interoperability Resources) server. The patient profile is very comprehensive, covering age, past medical history, current meds, and major diseases core component.
- **Tabbed Navigation:** In order to facilitate convenient use, the total health status of patients is divided into structured groups, making it easy to navigate through various aspects of patient care and information.
- **AI-Empowered Insights:** The integration of artificial intelligence enables real-time synthesis summaries on patient visits, medications dispensed, operations conducted, and clinical notes. This feature greatly improves the ability of healthcare personnel to make informed decisions in good time and helps boost patient outcomes.
- **Data Visualization:** Advanced visualization tools, such as time series diagrams, are used to depict the trends of various observations, such as physiological signs or results from laboratory examinations. This kind of pictorial representation is essential for understanding patient health tracks and making informed decisions at the clinical level.
- **Normalization and Error Handling:** Before data received from the FHIR server goes before the user, it goes through a normalization process. This step is

crucial for ensuring operations are conducted uniformly and avoiding misunderstandings, which may result from different data formats.

BACK TO LIST

Carli641 Ratke343 ID: 144

DEMOGRAPHICS ENCOUNTERS CONDITIONS OBSERVATIONS MEDICATIONS PROCEDURES DOCUMENTS

Documents + AI Summary

Search documents Filter by Author

ID	Date	Type	Status	Author	Actions
150	06/08/86, 05:46 AM	History and physical note Clinical Note	Completed	Dr. Raleigh478 Hauck852 COMMUNITY HEALTH CONNECTIONS, INC	View
156	06/19/88, 05:46 AM	History and physical note Clinical Note	Completed	Dr. Raleigh478 Hauck852 COMMUNITY HEALTH CONNECTIONS, INC	View
167	06/28/92, 05:46 AM	History and physical note Clinical Note	Completed	Dr. Raleigh478 Hauck852 COMMUNITY HEALTH CONNECTIONS, INC	View
179	05/06/09, 05:46 AM	History and physical note Clinical Note	Completed	Dr. Rayford811 Bashirian201 HENRY HEYWOOD MEMORIAL HOSPITAL	View
186	05/06/12, 05:46 AM	History and physical note Clinical Note	Completed	Dr. Raleigh478 Hauck852 COMMUNITY HEALTH CONNECTIONS, INC	View

Records per page: 5 1-5 of 57

AI Summary ×

Based on the provided medical notes, there are no specific cardiovascular concerns, diagnoses, or treatment recommendations directly mentioned in the patient's records.

Key observations regarding potential cardiovascular risk factors:

1. Patient is a 53-year-old female with no reported history of cardiovascular disease
2. Patient has several chronic conditions that could indirectly impact cardiovascular health:
 - Severe anxiety
 - Stress
 - Obesity (BMI 30+ in 2014)
3. Patient has never smoked, which is a positive factor for cardiovascular health
4. No cardiovascular-related medications or interventions are noted in the medical history

Potential cardiovascular risk areas to consider:

1. Chronic stress could potentially impact cardiovascular health
2. Obesity in 2014 suggests potential cardiovascular risk, though no recent BMI measurements are available
3. No documented cardiovascular screenings or assessments are mentioned in the provided records

Recommendation: While no immediate cardiovascular concerns are evident, routine cardiovascular health monitoring would be advisable given the patient's chronic stress and past obesity.

B. Back-end Architecture

The application backend serves as the brain of the entire system, implemented using Spring Boot as its core framework to centralize and automate various important tasks in terms of scheduling, data gathering, data association, and security management. [14] This advanced backend structure communicates effortlessly with the front-end interface through the use of Fast Healthcare Interoperability Resources (FHIR) and artificial intelligence (AI) models, thus allowing for the provision of actionable insights gleaned from intricate data sets. The backend encompasses several essential functions that contribute to the overall effectiveness and reliability of the system:

- **Data Processing and Routing:** The backend implements one unified REST API that allows routing requests from the front end. This API enables the exchange of many different types of data, such as information about medications, encounter data, and fees associated. This mechanism enhances the user experience as well as supports timely decision-making by ensuring the data flows smoothly and accurately from the front end to the backend.
- **FHIR Integration:** The backend implementation makes use of HAPI FHIR to ensure compliance with established healthcare interoperability standards. This library plays a vital role in receiving, processing, and normalizing patient data to adhere to a strict set of standards. Most importantly, this integration is key to improving the interoperability of healthcare data systems in order to enable better data sharing and use across the healthcare ecosystem.
- **AI Integration:** The backend architecture integrates cutting-edge artificial intelligence functionalities using four separate families of Gemini AI models that are specifically developed for varying facets of interactions with patients. These results are then aggregated and communicated back to the front end in a secure manner, allowing for the enhancement of care delivery through the power of data analysis.
- **User Management & Role-Based Control:** The backend has strong authentication mechanisms using OAuth2 and JSON Web Tokens (JWT) for security and access control. These technologies are intended to ensure secure access to the system while maintaining user privacy and data integrity. Additionally, future role-based access controls can be implemented for improved security.
- **Deployment and Scalability:** The backend is served through Docker containers, which provide a consistent, reliable service framework across a multitude of hardware configurations. Not only does this containerization create easier deployment processes, but it allows scalability, enabling the system to grow and evolve based on future needs and greater user demands. The backend can ensure high availability and performance to meet the needs of healthcare providers and their patients through the use of container technology.

Overall, the backend acts as a vital part of the system and utilizes several cutting-edge technologies and methodologies to offer healthcare IT solutions that are secure, efficient, and scalable, considering the constantly changing needs of the healthcare sector. It is responsible for improving the overall performance and reliability of the application through its multiple functions.

C. Security

Establishing a robust security framework is paramount in safeguarding sensitive healthcare data. In order to protect sensitive healthcare information, a strong security framework is of utmost importance. Various vital steps have been taken to reach that goal, forming an all-around protective system against threats to health information.

- **Authentication:** The OAuth2 protocol is used to provide authentication for secure and authenticated users. With this framework, each user gets access only after an extensive confirmation process, thus minimizing the chances of an unauthorized bearer. Alongside this, the JWT extension increases session security by transmitting the encoded information between parties in a secure manner. Implementing both of these technologies reinforces the authentication process and provides assurance to users that their personal health information is secure.
- **Role-based Access:** Access to sensitive health information has been restricted to qualified and authorized medical personnel, ensuring that only qualified personnel are allowed to access critical patient data. Nonetheless, it is important to realize that future versions of this system will focus on allowing role-based access control that goes beyond just medical providers. This transformation will allow patients and administrators to safely interact with the system and contribute to a more comprehensive approach to healthcare data management. This is likely to improve these access controls, offering permissions based on job roles, reducing the possibility of data breaches, and boosting security overall.
- **Data Normalization:** In this step, we perform data normalization, a key procedure used to standardize the information returned from the FHIR server. It is a crucial process to harmonize data within studies as it eliminates inconsistencies and misinterpretations that can affect the trustworthiness of data analyses. This uniformity of data allows us to make our analyses more accurate, thus trusting the end results based on that information. In its field of work, where the wrong analysis could lead to serious implications for patients' treatment and healthcare outcomes, diligence to data integrity is extremely critical.

In conclusion, the above measures are only the first of many steps that can be taken to ensure a secure environment in which sensitive healthcare data can ultimately be housed. As such, continuous improvement and adaptation to new security challenges will be critical to the ongoing integrity and confidentiality of health information in an ever-more digital environment.

D. Database Architecture

A dual-database architecture provides the first step in separating user authentication data and patient-related data. This design not only creates an environment where data can be secured, it also lends itself to operational efficiency, as disparate systems can be segregated and managed by these two fundamental aspects of information.

- PostgreSQL Database: The portal uses the PostgreSQL database for storing user profiles and session data in a secure way. This system is carefully developed to provide authentication methods and access modes to users on a one-to-one basis. The selected architecture involving an advanced relational database management system like PostgreSQL reflects a dedication to preserving excellence in data integrity and security, which is essential in handling sensitive user information.
- FHIR Server (HAPI FHIR): Adding a FHIR (Fast Healthcare Interoperability Resources) server with HAPI FHIR is a start towards improving healthcare data management. The FHIR server is designed to only provide compliant FHIR information and is based on the specs that can be found on HAPI for HL7. This server can handle requests that originate from web browsers and HTTP GET requests, which means that it returns only FHIR-compliant information about patients' health conditions. [11]

Moreover, the structure can dynamically manage data standardization and uniformity in the retrieval process. This modern architecture reduces the operational burden involved in data retrieval to a minimum and ensures that the system is scalable, secure, and efficient. The impact of such an architecture goes beyond offering basic functionality, and it provides an enriched user experience and augments the integrity of health information systems.

E. Tools and Technologies

The system integrates a suite of open-source tools that collectively enhance its functionality, scalability, and cost-effectiveness:

- Spring Boot: Spring Boot is a powerful framework that greatly improves the speed of building RESTful APIs and helps you build a rich backend infrastructure to support dozens of features. This helps developers build their applications faster by providing them with out-of-the-box features. This

framework has its own opinionated defaults and a large ecosystem that promotes best practices and microservices architectures. As a result of this systematic integration of best practices, Spring Boot not only facilitates rapid development cycles but also guarantees that the applications produced are performance-oriented, scalable, and consistent with modern software engineering practices. [14]

- Quasar Framework: One of the most dynamic and versatile frameworks for building responsive UIs specifically tailored for medical applications. Quasar has been able to achieve this remarkable feat by combining innovative technology with modern design principles, allowing developers to produce visually stunning, functionally rich, and accessible applications. This is a framework that allows for cross-platform development, enabling deployment across multiple devices and operating systems. This adaptability is especially beneficial in the highly specialized healthcare domain where user experience is key, and applications have to appeal to a wide range of users comprising healthcare professionals and patients. [13]
- Docker: Docker is a revolutionary containerization platform that provides consistency in deployments across multiple environments, including development, testing, staging, and production. Docker achieves this by packaging applications and their dependencies inside containers, reducing the possibility of inconsistencies that can occur when moving software through the various stages of the development lifecycle. The consistency of data is essential for preserving the integrity and reliability of applications, especially in the case of complex systems where multiple services are required to collaborate with each other smoothly. Additionally, Docker promotes better resource allocation, scalability, and quick iterations, allowing organizations to quickly adapt to market changes and emerging technologies. [15]
- HAPI FHIR: It is an open-source standard that facilitates the exchange of data in the healthcare domain. The core design principle of HAPI FHIR is to utilize the FHIR standard to achieve interoperability between otherwise heterogeneous healthcare systems, thereby improving the accessibility and usability of health information. The ability to make this possible through Data Exchange is especially relevant and impactful in healthcare, where timely and accurate data exchange can lead to improved patient care and clinical decisions. HAPI FHIR supports a more cohesive healthcare landscape by enabling seamless

connections between different healthcare systems, improving collaboration between providers, and ultimately enhancing patient care. [11]

- **Google Gemini AI Studio:** Such an option is not yet available for other data-collection tools, making this development a giant leap in clinical diagnostic capability. Describing underlying challenges that have faced healthcare professionals, especially in telling different antibiotics apart based on simple notes and ancillary objects with patients. This platform also provides sophisticated tools for classifying encounters by type to hospital infection control staff, pharmacy personnel, and medical techs, thus providing them the ability to classify encounters as medical treatment or non-medical context. This technology, Google Gemini AI Studio, helps improve patient safety and treatment outcomes through better diagnostic accuracy using artificial intelligence. [12]
- **Digital Ocean:** It is a reasonable cloud hosting platform best suited to commercial projects or open-source projects. It offers a powerful but cost-effective platform that empowers apps in the cloud by providing them with an infrastructure to use and a tech stack that they can build on in a way that keeps their costs in check. Developers easily use it at all experience levels due to its intuitive developer interface and thorough documentation. By leveraging capabilities offered by platforms such as Digital Ocean, healthcare institutions can deploy solutions at scale and speed without compromising the flexibility and efficiency needed to respond to the growing demands of the industry. [7]

These technologies, namely Spring Boot, Quasar Framework, Docker, HAPI FHIR, Google Gemini AI Studio, and Digital Ocean, together illustrate the transformative potential of modern software solutions in the realm of healthcare. Using these new capabilities, developers can innovate applications that help drive operational efficiencies as well as better care to patients.

F. Limitations

Although the project achieved its main objectives, some limitations were also noted:

- **Target Audience:** The system is currently for healthcare providers only. Patients or administrative staff would need different features and permissions for use.
- **Scope of Data:** The MVP pulls in just the 5,000 first records from the FHIR server, which is not very much for larger datasets. This limitation will be resolved on future releases with pagination.

- **Security Enhancements:** The system maintains secure authentication, though data is not encrypted. There is one big area for improvement in the future.

G. Future Enhancements

Depending on the success of the MVP, the following features will be introduced:

- **Pagination:** A good pagination would allow it to move through the system without saturating it with data.
- **Encryption:** There are plans to support encryption on sensitive data so stateful filtering will meet the specification level and availability.
- **Role-Based Access:** Adding different user roles will open the system to patients, administrators, and researchers.
- **Performance Enhancements:** Effectively minimize latency.

IV. IMPLEMENTATION AND TESTS

A. Implementation

The application was developed using a structured approach that was based on Agile methodologies, enabling iterative progress and flexibility during the project lifecycle. This process was effectively broken into four individual phases, popularly known as 'sprints,' each with specific goals and deliverables:

Sprint 1: The main goal of this first sprint was to set up a strong infrastructure for front-end and back-end services that help the application run. This foundational setup is important so that all the other phases of development can continue, as they would be working with the best faster and easier basis.

Sprint 2: This sprint mostly focused on model and API Development progress, and the application was also containerized under Docker, which is very important since it creates a portable and scalable way to bundle the application with all of its dependencies in one single container.

Sprint 3: This sprint focused on the UI component design and development, which can be dynamically loaded when required. This flexibility is essential for enhancing user experience. Simultaneously, work went into developing artificial intelligence (AI) models into complete software components that can provide actionable insights.

Sprint 4: The last development iteration of this series of sprints. In addition, there was thorough system deployment planning that supported the transition from development to production.

The deployment of a server-side engine for an FHIR platform was implemented in parallel with AI model algorithms that were containerized in Docker containers. Along with these emphases, front-end data visualizations and custom navigation improvements were implemented during this development phase. The decision taken to containerize every component with Docker played a crucial role here as it guaranteed a uniform testing area and simplified the deployment process to each environment.

B. Tests

The system has gone through extensive integration and unit testing before deployment. The thoroughness of this testing regime was critical to establishing that:

- The healthcare information remains accurate, reliable, and secure, as the data from the FHIR server is ingested, processed, and stored correctly.
- AI applications are integrated into the system's architecture and extract insights from existing data.
- Authentication and system-level access security protocols are properly implemented with a role-based access control paradigm that protects sensitive information while safeguarding user privacy.
- The application was hosted on the Digital Ocean cloud, and it took advantage of Docker Compose to facilitate deployment. It highly contributed to making the deployment strategy deploy backend, frontend, and the database as a service.
- Utilizing Docker Compose, we were able to simplify the management and orchestration of the individual services, as well as ensure that the deployment process aligned with best practices in both software development and operations.

V. CONCLUSION

This paper shows the power of healthcare through open-source technologies and how they can decrease costs, encourage collaboration, and retain flexibility. The proposed system demystifies healthcare innovation based on health organizations by means of frameworks and tools (Spring Boot, Quasar, Docker, and HAPI FHIR), enabling a fast transformation of obsolete trends without the economic or technical awakening usually seen in the literature. Such care inside vibrant communities around these open-source solutions finds a way to maintain system relevance, working in response to the changing requirements of healthcare providers, and so we are embedded in a spiral of progress and improvement.

On the other hand, Artificial intelligence works well with Fast Healthcare Interoperability Resources (FHIR) standards to ensure that complex healthcare data can be consumed seamlessly, making it easier for providers to help improve patient care. By doing so, this system not only eliminates cumbersome processes that have been traditionally associated with healthcare data management but also increases data availability and enables providers with access to real-time data to make informed and data-driven decisions.

While the project achieves its fundamental objectives, it highlights areas for possible improvements, such as the application of encryption protocols for sensitive data, the inclusion of role-based access controls, and the optimization

of processing capabilities for larger data sets. These limitations provide boundaries for future development to keep the application relevant and responsive to the ever-changing medical landscape.

In summary, this paper shows great developments in the field of health informatics. It explores how, by tackling modern issues while establishing the foundations for advancements to come, leveraging open-source tech, artificial intelligence, and FHIR standards can collectively transform care provision, uplift clinical outcomes, and progress the entire healthcare informatics domain.

REFERENCES

- [1] [1] L-M. Junttila, Peltonen, and S. Salanterä, "Front-Line Physicians' Satisfaction with Information Systems in Hospitals . in Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth," *Studies in Health Technology and Informatics*, vol. 247, no. 0, pp. 865-869, 2018, Available: <https://core.ac.uk/download/287761701.pdf>
- [2] [2] H. Larsen and P. Spagnoletti, "Secure data exchange in healthcare infrastructures A case study of Norway and Denmark," 2024. Accessed: Dec. 07, 2024. [Online]. Available: <https://core.ac.uk/download/616581662.pdf>
- [3] [3] N. Chen, "Smart Healthcare solutions in China and Europe, an international business perspective," 2023. Accessed: Sep. 06, 2024. [Online]. Available: <https://core.ac.uk/download/491249162.pdf>
- [4] [4] Rishabh Kapoor, "A Learning Health System for Radiation Oncology," *Core.ac.uk*, vol. 1, 2023, doi: oai:scholarscompass.vcu.edu:etd-8575.
- [5] [5] A. Koumpis, K. Kalafatakis, F. Limani, and M. Bregonzio, "Advancing healthcare through data: the BETTER project's vision for distributed analytics," 2024. Accessed: Dec. 07, 2024. [Online]. Available: <https://core.ac.uk/download/621428829.pdf>
- [6] [6] J. Wilser, "Reducing Doctors' Paperwork," *TIME*, Oct. 30, 2024. <https://time.com/7094840/abridge/> (accessed Dec. 07, 2024).
- [7] [7] DigitalOcean, "Product Docs Home | DigitalOcean Documentation," *Digitalocean.com*, 2024. <https://docs.digitalocean.com/products/>
- [8] [8] B. Blobel and F. Oemig, "Solving the Modeling Dilemma as a Foundation for Interoperability," *European Journal for Biomedical Informatics*, vol. 14, no. 3, 2018, doi: <https://doi.org/10.24105/ejbi.2018.14.3.2>.
- [9] [9] H. Mag, "Open Source Will Help Drive EHR Costs Down," *Hcinnovationgroup.com*, Aug. 27, 2009. <https://www.hcinnovationgroup.com/policy-value-based-care/health-it-leadership/article/13000903/open-source-will-help-drive-ehr-costs-down> (accessed Dec. 07, 2024).
- [10] [10] Ordina Belgium, "Introduction to HAPI FHIR - Youri Vermeir," *Ordina-jworks.io*, 2021. <https://blog.ordina-jworks.io/ehealth/2021/02/23/hapi-fhir.html> (accessed Dec. 08, 2024).
- [11] [11] "Index - FHIR v4.0.1," *www.hl7.org*. <https://www.hl7.org/fhir/>
- [12] [12] Gemini, "Google AI Studio quickstart," *Google AI for Developers*, 2024. <https://ai.google.dev/gemini-api/docs/ai-studio-quickstart>
- [13] [13] Quasar, "Quasar Framework - Build high-performance VueJS user interfaces in record time," *Quasar Framework*. <https://quasar.dev/>
- [14] [14] Spring, "Spring Projects," *Spring.io*, 2019. <https://spring.io/projects/spring-boot>
- [15] [15] Docker, "Docker Desktop overview," *Docker Documentation*, Dec. 23, 2021. <https://docs.docker.com/desktop/>

Yapay Zeka Destekli Uçtan Uca İşe Alım Uygulaması

Ferdi Sönmez
Yazılım Mühendisliği
İstanbul Aydın Üniversitesi
İstanbul, Türkiye
ferdisonmez@aydin.edu.tr

Ercan Öge
ABMYOO İnsan Kaynakları Bölümü
İstanbul Aydın Üniversitesi
İstanbul, Türkiye
eoge@aydin.edu.tr

Özet—İşletmelerin varlıklarını devam ettirmeleri ve rekabet üstünlüğünü sağlamaları ancak nitelikli insan kaynağı ile gerçekleştirilmektedir. Nitelikli çalışanların işletmeye kazandırılması işletmelerin öncelikli konusu ve aynı zamanda da sıkıntı yaşadıkları konuların başında gelmektedir. Dolayısıyla işe alım sürecinin önemli aşamalarından biri olan aday araştırma ve bulma sürecinin doğru yönetilmesi işletmeler için önem arz etmektedir. Yapay zekâ tüm işletme fonksiyonlarında olduğu gibi İnsan Kaynakları (İK) süreçlerinde de ciddi değişimler yaratmakta ve yaratmaya devam edecektir. Bu çalışmanın temel amacı da İK süreçlerinden biri olan işe alım sürecinde geliştirilecek olan algoritma ve modeller (yapay zeka) ile toplanan özgeçmişleri analiz ederek boş pozisyona dair en uygun ve güncel yetkinlikleri uyumlaştırmak ve bu yetkinlikleri ağırlıklandırmak, başvuruların özgeçmişleri puanlamak ve sıralamaktır. Bununla birlikte, aday ve kurum arasındaki operasyonel diğer süreçlerin yönetilmesini sağlayan sanal asistanı; işe yerleştirme noktasında özlük belge yönetimi işini otomatize eden ve işten ayrılma ihtimali tespiti de yapan uçtan uca süreç yönetimi ve karar destek sistemini; geliştirmek ve işe alım sürecine yönelik sorunlara çeşitli çözümler getirmek olup, bu ilişkinin nasıl modellenebileceğini ortaya çıkarmaktır. Bu çalışmanın kaynağını oluşturan ve açık kaynak kod desteğini de alınarak geliştirilen proje çalışması ile uçtan uca alım süresinin ortalama en az %35 azaltılması, işe alım maliyetinin (zaman) şirket ölçeğine göre en az %30 azaltılması, uygun aday bulma süresinin en az %25 azaltılması ve adayların işe başvuru işleminin geleneksel yöntemlere göre en az %40 hızlandırılması sağlanmıştır.

Anahtar Kelimeler— İKY, yapay zeka, açık kaynak kod, sanal işe alım, işe alım

Abstract—The survival of businesses and their competitive advantage can only be achieved with qualified human resources. Bringing qualified employees to the business is a priority for businesses and also one of the issues they experience difficulties with. Therefore, the correct management of the candidate research and finding process, which is one of the important stages of the recruitment process, is important for businesses. Artificial intelligence creates and will continue to create serious changes in Human Resources (HR) processes as well as in all business functions. The main purpose of this study is to analyze the resumes collected with the algorithm and models (artificial intelligence) to be developed in the recruitment process, which is one of the HR processes, to harmonize the most appropriate and up-to-date competencies for the vacant position and to weight these competencies, to score and rank the resumes of the applications. In addition, the virtual assistant that enables the management of other operational processes between the candidate and the institution; the end-to-end process management and decision support system that automates the personnel document management process at the point of

placement and also detects the possibility of leaving the job; and to develop various solutions to the problems related to the recruitment process and to reveal how this relationship can be modeled. With the project work that forms the source of this study and was developed with open source code support, the end-to-end recruitment time was reduced by at least 35% on average, the recruitment cost (time) was reduced by at least 30% depending on the company scale, the time to find a suitable candidate was reduced by at least 25% and the job application process of the candidates was accelerated by at least 40% compared to traditional method..

Keywords— HRM, artificial intelligence, open source, virtual recruitment, recruitment

I. GİRİŞ

İşletmelerde ihtiyaç duyulan pozisyonlar için aday bulma; sadece işgören seçimi için değil aynı zamanda diğer insan kaynakları yönetimi (İKY) işlevleri açısından da önem arz etmektedir. İnsan kaynaklarının (İK) kalitesi, aday bulmadaki başarısına da bağlıdır. Uygun özelliklerde ve gereken sayıda adayın bulunamaması, bazı boş pozisyonların doldurulamamasına sebep olacağı gibi pozisyonun gerektirdiği niteliklere yeterince sahip olmayan ya da fazla nitelikli kişilerin işe alınmasına da sebep olabilir [1]. Bu durum da işletme açısından maliyet, aynı zamanda işlerin devamlılığı açısından önemli sorunlar ve sonuçlar doğurması beklenmelidir.

Jennings ve Hauhton'un [2] belirttikleri üzere iş yaşamında başarılı olmak adına başarılı olmak için artık büyük olmak yeterli görülmemektedir. Bununla birlikte; hızlı düşünmek, hızlı kararlar almak, rakiplere göre daha hızlı harekete geçmek başarının yeni ölçütleri olarak ön plana çıkmaktadır. Hızlı düşünmek, hızlı kararlar almak, rakiplere göre daha hızlı harekete geçmek tüm işletme süreçlerinde söz konusudur. İKY'nin önemli süreçlerinden biri olan işe alım açısından da belirtilen hususlar hem rekabet açısından hem de maliyetler açısından işletmelere önemli kazanımlar sağlamaktadır.

Son yıllarda teknolojiye ortaya çıkan baş döndürücü gelişmeler, işletmeleri her alanda arayışa itmiştir. İşletmeler de bu gelişmeleri süreçlerine başarılı şekilde bütünleştirmeye çalışmaktadır. İşe alım sürecinin önemli bir aşaması olan özellikle aday araştırma ve bulma süreci de bunlardan birisidir [3]. İşletmelerin İK Birimleri bu konuda teknolojinin nimetlerinden yararlanmaya başlamışlardır. Bu amaçla Birimleri ya kendi bünyesinde oluşturdukları bir ekip (Teknoloji ve İKY Birimindeki çalışanlardan oluşan bir ekip) aday araştırma ve başvurularını sağlama sürecini yönetmeye çalışmakta ya da bu süreçlerle ilgili dışarıdan profesyonel bir destek almaktadır.

Yapay zeka konusu ile ilgili son yıllarda gelişmeler ışığında literatürde [4] [5] yapaya zeka ile işe alım süreci

ilişkinine yönelik çeşitli çalışmalar söz konusudur ve artan bir hızla devam etmektedir. Ancak konuya ilişkin yazılım modelleri sınırlı olup [5], bu modellerin sınırlılığı çerçevesinde bu çalışma (proje) ile hem literatüre hem de iş yaşamına bu yönde katkı sağlanmaya çalışılmıştır. Çalışmanın ikinci bölümünde **ilgili literatür**, üçüncü bölümünde veri toplama, dördüncü bölümünde **bulgular ve beşinci bölümünde sonuç yer almaktadır.**

II. İLGİLİ LİTERATÜR

İşe alım sürecinde zaman çok önemli bir husus olup, bununla ilgili üç tane önemli alt konu bulunmaktadır. Birinci konu olan cevap verme süresi, nitelikli aday havuzu oluşturmak ve görüşme için bu adayları mülakata çağırarak için geçen süredir.

A. İşe alım sürecinde zaman

Cevap verme süresi, işletme için kritik olarak kabul edilen pozisyonların doldurulmasında ve özellikle de işgücü pazarının dar olduğu dönemlerde büyük önem arz etmektedir. Böylelikle, işe alım sürecinde zaman, insan kaynağı temininde önemli bir ölçüt olarak ele alınmaktadır [6]. Zamanla ilgili olan ikinci konu ise pozisyonun ne kadar zamanda doldurulacağı yani pozisyonun doldurulma süresidir [7][8]. Zamanla ilgili üçüncü konu ise, işe yeni kabul edilen kişinin işe fiilen başlama zamanı olan işe başlama zamanıdır [6][7]. Bu aşamada İK departmanının kontrolü fazla değildir. Cevap verme süresi, İK departmanına arzu edilen eleman talebine ilişkin resmi bir belgenin ulaşmasından, mülakat gününe kadar geçen süredir ve departmanın ne kadar zamanda temin sistemini çalıştırabildiğinin önemli bir göstergesidir. İstenilen nitelikteki elemanların işletmeye çekilebildiği süre olgusu İK departmanı kadar işletme için de önemlidir [9]. Bir yönetici (temin veya kadrolama), değerlendirme sistemini geliştirmek için boş olan pozisyonun doldurulma süresi ile işe başlama zamanı değerlerini düzeltmeye yönelik bilgi bankası da oluşturmayı tercih edebilir [9].

Cevap verme süresi [6], boş olan pozisyonların doldurulma süresi ve başlama zamanı hakkında kayıtların saklanması hem yöneticilere, hem de İK departmanına süreçle ilgili kıymetli bilgi sağlayacaktır. Yöneticiler gerçekleştirdikleri işi ve performanslarını sayısal değerlerle destekleyebilecekleri gibi, elde bulunan bu bilgiler yardımıyla tepe yönetimden süreçlerin hızlandırılması veya işbirliği elde etmek için destek isteyebileceklerdir. Dönemsel olarak, temin sürecinin etkinliğini gözlemlemek ve karşılaştırmalar yapmak çok önemlidir [9].

İK uzmanları mevcut verileri dikkatlice incelemese ciddi organizasyonel problemleri fark edemezler. İş duyurularına talep oranının hesaplanmasının dışında, İK departman tarafından izlenmesi gereken bir ölçüm de duyurulan işlerin ne kadarına cevap verildiğidir. Bu ölçüme, elle veya otomatik olarak yapılan basit bir çizelge ile ulaşılabilir. Tablo I buna bir örnek olarak verilebilir [9].

TABLO I: İŞ DUYURULARINA TALEP ORANI

İş duyurularına talep oranı	
Duyurulan işlerin cevaplanma oranı	$IDCO = \frac{CDIS}{DI}$
Burada:	
CDIS =	Cevaplanan iş duyurularının sayısı
DI =	Duyurulan işlerin sayısı

Temin sürecinin etkililiğinin artırılmasında değerlendirme önemli olup, aşağıdaki faaliyetlerin dönemsel olarak gözden geçirilmesi ve gerekiyorsa düzeltici önlemlerin alınması önemlidir. Bunlar:

- **Özgeçmiş Madenciliği:** İşletme ile uyumunun yüksek olacağı tahmin edilen özgeçmişlerin işletme bilgi sisteminden çıkarımını sağlar,
- **Aday Takip Sistemi:** Adayı temin sürecinde itibaren performans değerlendirme sonuçlarına kadar ele alır,
- **İşletmeye özgü web sitesi:** Boş pozisyonların duyurulduğu web sitesi/sayfası,
- **İç hareketlilik:** İşletmedeki gelecek vaat eden çalışanların takip edildiği ve olası adayların karşılaştırıldığı sistem, • **Gerçekçi iş görüşmeleri:** İşletme ve işler konusunda tüm gerçeklerin adaylarla paylaşıldığı temin sistemi,
- **Adaylara zamanında cevap verilmesi:** Adayların işletmeden zamanında cevap aldığı sistem, olarak tanımlanabilir [9].

B. Verimlilik Açısından İnsan Kaynakları Seçim Sürecinin Değerlendirilmesi

Seçim sürecinin kalitesinin belirlenmesi kısa vadede çok zordur. Bu olumsuz durum uzun vadede performansı ve kıdemi de etkileyebilir. Seçim sürecinde ilerleme gösteren aday(lara) dair iki önemli ölçüm kriteri işlemlerin hızı ve aşamaların maliyetidir.

Seçim sürecinin verimliliği, seçilen adayların sayısı ve işe alınmalarına odaklanmaktadır. Burada ölçüm kriteri belirli bir zaman diliminde ilerleyen çalışanların sayısıdır. İşe alma sürecinin hacmine de bu sayı ile ulaşılabilir [9].

C. Maliyetler Açısından İnsan Kaynakları Seçim Sürecinin Değerlendirilmesi

Çeşitli seçim öğelerinin etkinliğinin karşılaştırılmasında kullanılan önemli bir maliyet hesabı da seçim sürecine ilişkin maliyetler olarak ortaya çıkmaktadır. Maliyetlerin belirlenmesinde:

- Yeni işe kabul edilenlerin işe alma maliyeti (işe alma maliyeti)
- Belirli bir yöntemle seçilenleri işe alma maliyeti (değerleme merkezi, testler, sadece görüşme)
- Belirli bir gruba göre maliyetlerin ortaya konması (yönetici - yönetici olmayan, kapsam içi - kapsam dışı gibi) gibi ölçümlere başvurulmaktadır.

Toplam maliyetler, kadrolama elemanlarının maaşları, adayların ve diğerlerinin seyahat masrafları, temin maliyetleri ve diğer direkt maliyetler olarak sıralanabilir [9].

III. VERİ TOPLAMA

Proje kapsamında geliştirilen üründe Selenium Framework u tercih edilerek [10] Açık kaynaklı bir programlama dili olması ve diğer açık kaynaklı bir dillerinin

tüm özelliklerine sahip olması nedeniyle tercih edilen Python programlama dilinde öncelikle LinkedIn üzerinde çalışacak bir script yazılmıştır. Bu scrapper LinkedIn Jobs üzerindeki verilen “title” ve “seniority” ile arama yaparak ilan tarihi sıralamasına göre ilanın “seniority”, “title”, “industry” “work type” ve tanımını alarak ham bir şekilde MongoDB üzerine kaydetmektedir.

Tanım üzerinde geliştirilen NER Modelleri (Named Entity Recognition - İsimlendirilmiş Varlık Tespiti) [11] çalıştırılarak eğitim, beceri, sertifika, dil alanları yapısal biçimde MongoDB üzerine ayrıca kaydedilmektedir.

Bu yöntemle 107 pozisyon için toplam 30 bin iş ilanı toplanmıştır.

A. NER MODELLERİ

Toplanan iş ilanı metinleri ve özgeçmişler; etiketlenerek NER modelleri için veri setleri oluşturulmuştur. 5 farklı NER modeli eğitilmiştir. Bu modeller ve varlıkları Tablo II’de gösterilmektedir.

TABLO II: MODELLER VE VARLIKLARI

Varlık	Eğitim	Deneyim	Genel	Beceri	Sertifika
Kurum	x				
Anadal	x				
Derece	x				
GPA	x				
Unvan		x			
Şirket		x			
Tarih			x		
Şehir			x		
Ülke			x		
Dil			x		
Beceri				x	
Sertifika					x

B. Sıklık ve Sıralama Modeli

Sistemde mevcut olan kıdem (7 adet) ve pozisyonlara (107) eşlenen LinkedIn üzerinden alınan iş ilanlarındaki eğitim, beceri, sertifika, dil ve tecrübe bilgileri veri seti olarak kullanılmıştır. Bu iş ilanlarındaki geçen anahtar değerlerin iş ilanı metinlerindeki pozisyonları yani yukarıdan aşağıya doğru yer alma sıralama olarak belirlenmiştir. Aynı anahtar kelimelerin sıklık değeri, kaç iş ilanında geçtiği, belirlenmiştir. TF-IDF mantığına (Term Frequency — Inverse Document Frequency, Terim Frekansı-Ters Belge Frekansı) benzer anlaşılır ve açıklanabilir bir model oluşturulmuştur [12]. Sıklık ve sıralama özelliklerinin ağırlıkları başlangıçta %50 ve %50 şeklinde belirlenmiştir. Bu ağırlıklar sistem kullanıcılarının hareketlerinden oluşacak geri dönüşlerle optimize edilecek şekilde planlanmıştır. Bu iki özellik üzerinden Kıdem x Pozisyon kombinasyonu (749 kombinasyon) için sıralama ve sıklık üzerinden anahtar değerlere 0-1 kapalı aralığında değerler atanmıştır. Değerler belirli bir eşik değerinin üstünde olduğu durumlarda o kombinasyon için öneri olarak sunulmuştur. Using the Template

C. Birliktelik Kuralı Modeli

Birliktelik Kuralı görece büyük veri setlerindeki ilişkisellikleri anlamak için kullanılan bir yöntemdir [13]. E-ticaret ve satışa yönelik alanlarda bu yöntemi sıklıkla kendi problemimize uygulayarak sonuçları gözlenmiştir. Sistemdeki iş ilanlarından faydalanarak hangi kıdem ve pozisyonların hangi anahtar kelimelerle daha sık bir arada bulunduğu

belirlenmiştir. Bu modelin çıktısı olarak 3 değer oluşturulmuştur:

-Support (Destek): Anahtar değer sıklık oranını, tüm veri seti üzerinden oluşturan değer.

-Confidence (Güven): Bu değer koşullu olasılık formülasyonu üzerinden, anahtar kelimenin o iş ilanında bulunma olasılığının hesaplandığı değerdir. 0 ve 1 arasında bir değer alır. Değer 1'e ne kadar yakınsa o ilişki o kadar kuvvetlidir.

- Lift (Kaldıraç): Bu değer üstteki iki değer birbirine olan oranı olarak hesaplanmaktadır. Bu değer 1’ den yüksek olması anahtar değer o iş ilanı içerisinde yer alma olasılığına pozitif olarak etki ettiğini göstermektedir. Lift değeri ne kadar yüksekse birliktelik o kadar güçlüdür. Bu değerler arasından karar vermek için öncelikle Lift değeri kullanılmıştır. Belirlenen bir eşik değerinden yüksek Lift değerine sahip olan anahtar değerler, o kombinasyon için önerilmiştir.

Model çıktılarının başarısını ölçmek için, 2 bağımsız IT recruiter (BT işe alımcı) tarafından hazırlanan önerilerle model önerilerinin karşılaştırılmıştır.

IV. BULGULAR

Proje Bu çalışmanın temelini oluşturan proje BİS Çözüm firmasının TÜBİTAK TEYDEB Ar-Ge Destek Programı kapsamında desteklenen ‘Yapay Zeka Destekli Uçtan Uca İşe Alım Uygulaması’ başlıklı projesidir. Aşağıdaki başlıklar projenin başarı ölçütlerini ve sonuçlarını ifade etmektedir.

A. Uçtan Uca İşe Alım Sürecinde Hız Artışı

Uçtan uca işe alım sürecinde hız artışı ölçütünde dayanak olarak ele alınan ve proje öncesi yapılan araştırmalar sonucunda ulaşılan uçtan uca alım süresinin ortalama 36 gün olduğu tespit edilmiştir. Bu sürenin en az %35 azaltılması sonucu ortalama 25 günün altına indirilmesi hedeflenmiştir. Böylelikle sürece dair aşağıdaki süre tanımları;

- İlan hazırlanması ve yayınlanması 1 gün
- Başvuruların toplanması ve aday skorlarına göre sıralamanın yapılması 6 gün
- İlk mülakat aşaması 7 gün
- İkinci mülakat aşaması 7 gün
- Teklif ve işe alım 2 gün

olarak hedeflere ulaşılmaması söz konusu olmuştur..

B. İşe alım maliyetinde azalma

İşe alım maliyetinde azalma başarı ölçütünde, proje öncesi yapılan araştırmalar sonucunda işe alım maliyeti (zaman) şirket ölçeğine göre değişmekte olup ortalama \$ 4000 - \$ 20000 arası olduğu tespit edilmiştir. Bu maliyetin en az %30 azaltılması sonucu ortalama \$ 2800 - \$ 14000 indirilmesi hedeflenmiştir.

İş ilanındaki pozisyona dair önerilen yetkinliklerin uygun ve güncel olması gerekmektedir. Bu başarı ölçütünde hedeflenen başarı en az %95 olarak belirlenmiştir. Bu hedefin doğru ölçümü için 2 bağımsız işe alımcıdan aşağıda belirtilen pozisyonlar için iş ilanı gereksinimleri hazırlanması talep edilmiştir. Bu isterler sistem işleyişine uygun olması adına **Eğitim, Dil, Beceri, Tecrübe ve Sertifika** kategorilerinde aşağıda listelenen isterler hazırlanmıştır:

- Data Scientist
- Ios Developer
- Devops Engineer
- ETL Developer
- Machine Learning Engineer
- Frontend Developer
- Flutter Developer

Sistem üzerinden alınan isterler ile işe alımcıların hazırladığı isterler için karşılaştırma tablosu (Tablo III) ve ekran görüntüleri (Fig. 1) aşağıdaki gibidir.

TABLE III. İSTERLER KARŞILAŞTIRMA TABLOSU

	Recruiter 1 Accuracy	Recruiter 2 Accuracy
	İşe Alımcı 1	İşe Alımcı 2
Oran	%96,10	%94,80

Kategori	Recruiter 1	Recruiter 2
Senior Data Scientist	62,5	61,7
Senior Data Scientist	54,2	51,1
Senior Data Scientist	49,3	48
Senior Data Scientist	47,7	35,9

Fig. 1. İsterler karşılaştırması

C. Uygun Aday Bulma Süresinde Azalma

Proje öncesi yapılan araştırmalar sonucunda uçtan uca alım süresinin ortalama 36 gün olduğu tespit edilmiştir. Bu sürenin en az %25 azaltılması sonucu ortalama 25 günün altına indirilmesi hedeflenmiştir. Uygun aday bulma hedefine ulaşmak için en önemli adım olan skora göre sıralama stratejisinin ne kadar iyi çalıştığını görmek için aşağıda belirtilen iş ilanına (Fig. 2) başvuran 8 adayın CV skorlarına göre sıralaması hem skora göre model hem de 2 bağımsız işe alımcı tarafından yapılmıştır.

Fig. 2. İş İlanı

CV leri sistem tarafından puanlanmış ve sıralanmış hali Fig. 3'deki gibidir.

CV No	Ad Soyad	Puan
CV-1	B... K...	60
CV-2	D... B...	67
CV-3	A... D...	85
CV-4	E... E...	55
CV-5	M... H...	90
CV-6	M... Y...	65
CV-7	M... M... A...	70
CV-8	S... D...	60

Fig. 3. Sistem tarafından puanlanmış ve sıralanmış CV ler

Bağımsız işe alımcılar tarafından hazırlanan sıralama ise Tablo IV'deki gibidir. Burada, adayların isimleri KVKK (GDPR) kapsamında maskeleymiş, yerine CV-1 başlayan bir kodlama ile ifade edilmiştir (Tablo IV).

TABLE IV: BAĞIMSIZ İŞE ALIMCILAR TARAFINDAN HAZIRLANAN SIRALAMA

CV	İşe Alımcı 1	İşe Alımcı 2
B... K... (CV-1)	60	55
D... B... (CV-2)	67	70
A... D... (CV-3)	85	90
E... E... (CV-4)	55	55
M... H... (CV-5)	90	85
M... Y... (CV-6)	65	70
M... M... A... (CV-7)	65	70
S... D... (CV-8)	60	70

Sistemin puanlaması ile bağımsız işe alımcıların puanlaması arasında fark olsa da sıralama noktasında oldukça benzer bir sonuç ortaya konulduğu görülmektedir.

D. İşe Başvuru Adımının Hızlanması

Bu başarı ölçütünde adayların başvuru işleminin geleneksel yöntemlere göre en az %40 hızlandırılması hedeflenmiştir. Bu hedefin doğruluğunu kontrol etmek adına 8 farklı CV ile projedeki başvuru modülünden başvuru yapılarak zaman etüdü yapılmıştır. Aynı zaman etüdü farklı kaynaklardan da başvuru yapılarak ölçülmüştür. Başvurularda kullanılan CV listesi, süre karşılaştırmaları ve iyileşme oranları hakkında bilgi aşağıdaki gibidir.

Burada da adayların isimleri KVKK (GDPR) kapsamında maskeleymiş, yerine CV-11 başlayan bir kodlama ile ifade edilmiştir. Tabloda görüldüğü üzere ortalama bir başvuru 37,6 saniye sürmektedir. Sürelerin tespitinde kullanılan CV'ler kişisel bilgi içerdiğinden bu çalışmada ek olarak paylaşılmamıştır (TABLO V).

Tablo V: CV listesi, Süre Karşılaştırmaları

CV	Type (Tür)	Personai Parsing Duration (Ayrıştırma Süresi)	Personai Correction Duration (Düzeltilme Süresi)	Personai Total Duration (Toplam Süre)
CV-11	LinkedIn Format	22,40	18,7	41,10
CV-12	Custom Format	10,24	21,1	31,34
CV-13	LinkedIn Format	11,51	17,2	28,71
CV-14	Kariyer.net Format	16,00	19,2	35,20

CV-15	Custom Format	45,70	6,1	51,80
CV-16	LinkedIn Format	24,42	19,4	43,82
CV-17	Custom Format	18,29	11	29,29
CV-18	LinkedIn Format	30,15	9,4	39,55
Average Duration		22,34	15,26	37,60

V. SONUÇ

Bu bildirinin temelini oluşturan çalışma BİS Çözüm firmasının TÜBİTAK TEYDEB Ar-Ge Destek Programı kapsamında desteklenen ‘Yapay Zeka Destekli Uçtan Uca İşe Alım Uygulaması’ başlıklı projesidir. Çalışma, işe alım süreci ile yapay zeka konusunu ele alan bir model önerisinden veya uygulamasından oluşmaktadır. Proje kapsamında ortaya çıkarılan model sonucunda elde edilecek ‘işe alım süreç problemlerine yönelik doğal makine/derin öğrenme tabanlı uçtan uca sürecini otomatize eden karar destek sistemi’ sayesinde işletmelerde nitelikli insan kaynağı vasıtası ile yurt içi ve yurtdışı pazarlarda rekabet gücünün artması ve bu amaca yönelik hazırlanan söz konusu yazılımın ithal ikamesi etkisini azaltarak döviz kazandırıcı nitelik taşıması beklenmektedir. Elde edilen bulgular ve sonuçlar da bunu destekler niteliktedir.

KAYNAKLAR

- [1] [Uyargil, C. (2023). İnsan kaynakları Yönetimi, 10. Baskı, Beta Yayınları, 2023.
- [2] Jennings, J., & Haughton, L. (2002). Its not the Big that Eat the Small... Its the Fast that Eat the Slow. New York: HarperCollins Publishers.
- [3] Garg, S., Sinha, S., Kar, A. K., & Mani, M. (2022). A review of machine learning applications in human resource management. International Journal of Productivity and Performance Management, 71(5), 1590-1610.
- [4] Delecraz, S., Eltarr, L., Becuwe, M., Bouxin, H., Boutin, N., & Oullier, O. (2022). Responsible Artificial Intelligence in Human Resources Technology: An innovative inclusive and fair by design matching algorithm for job recruitment purposes. Journal of Responsible Technology, 11, 100041.
- [5] Ore, O., & Sposato, M. (2022). Opportunities and risks of artificial intelligence in recruitment and selection. International Journal of Organizational Analysis, 30(6), 1771-1782.
- [6] Majumder, S., & Mondal, A. (2021). Are chatbots really useful for human resource management?. International Journal of Speech Technology, 24(4), 969-977.
- [7] Durai, S., Krishnaveni, K., & Manoharan, G. (2022, May). Leveraging HR metrics for effective recruitment & selection process in IT industries in Chennai and Coimbatore, Tamil Nadu. In AIP Conference Proceedings (Vol. 2418, No. 1). AIP Publishing.
- [8] Wandhe, D. P. (2020). HR Analytics: A tool for strategic approach to HR productivity. Available at SSRN 3700502.
- [9] Tüzüner, V.L. (2011). İnsan Kaynakları Yönetimi Faaliyetlerinde Ölçme ve Değerlendirme, 2. Baskı, Beta Yay.
- [10] Horalek, J., Urbanik, P., & Sobeslav, V. (2022, October). Automated Tests Using Selenium Framework. In International Conference on Nature of Computation and Communication (pp. 23-40). Cham: Springer Nature Switzerland.
- [11] Li, J., Sun, A., Han, J., & Li, C. (2020). A survey on deep learning for named entity recognition. IEEE transactions on knowledge and data engineering, 34(1), 50-70.
- [12] Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. Plos one, 16(8), e0254937.
- [13] Karasu, B., & Doğan, O. (2020). BİRLİKTELİK KURALLARI MADENCİLİĞİ VE İLGİNÇLİK ÖLÇÜMLERİ: BİR VAKA ÇALIŞMASI. Journal of Business in The Digital Age, 3(2), 94-107.
- [14] Bibi, M., Abbasi, W. A., Aziz, W., Khalil, S., Uddin, M., Iwendi, C., & Gadekallu, T. R. (2022). A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis. Pattern Recognition Letters, 158, 80-86.

A MATLAB App for Decomposing the Even and Odd Parts of a Transfer Function

Ramazan Menak
Electrical-Electronic Engineering
Siirt University
Siirt, Türkiye
ramazanmenak@siirt.edu.tr

Nusret Tan
Electrical-Electronic Engineering
İnönü University
Malatya, Türkiye
nusret.tan@inonu.edu.tr

Abstract—In this paper, a MATLAB App program is presented that decomposes the numerator and denominator polynomials of transfer functions into their odd and even components. In control engineering, transfer functions are essential tools for modeling system dynamics, and decomposing these functions into their odd and even components is a key step in system analysis. The developed MATLAB App allows users to enter the numerator and denominator polynomials of transfer functions efficiently, automatically separating them into odd and even parts. This functionality enables researchers to examine system characteristics in more detail and perform essential analyses with ease. The MATLAB App is designed to reduce time and errors, making complex mathematical operations accessible through a user-friendly layout. This App aims to improve researchers' productivity and increase the efficiency of their analysis processes.

Keywords—Even and odd parts, Matlab App, Matlab GUI, Stability boundary locus

I. INTRODUCTION

In engineering, particularly in the field of control systems, system modeling and analysis form the cornerstone of design processes. Effectively managing these processes requires the ability to perform complex mathematical computations quickly and accurately. In recent years, various software tools have been developed to accelerate these calculations and make them more user-friendly. Among these tools, MATLAB has found widespread use in engineering applications. With its powerful mathematical capabilities and user-friendly interfaces, MATLAB is frequently preferred in engineering tasks. The software's Graphical User Interface (GUI) tools enable users to perform complex calculations and simulations more easily and efficiently, making it an invaluable resource for engineers and researchers in control systems. In recent years, MathWorks has discontinued support for the GUIDE environment and introduced a new tool, MATLAB App Designer [1], as its replacement. As a result, numerous programs continue to be developed with the same objective of facilitating researchers' operations and enhancing the efficiency of their workflows.

The teaching of control systems has been significantly enhanced by the development of interactive tools and software platforms. For instance, Andreatos and Zagorianos introduced a MATLAB-based user-friendly graphical user interface (GUI) designed specifically for teaching Automatic Control Systems. Their work focused on creating an intuitive platform that facilitates the simulation and comprehension of control systems, offering educators and learners an effective tool to enhance the learning process [2]. Interactive learning tools have proven to be essential in enhancing the comprehension of frequency response methods in control systems. Mitchell [3] developed two graphical user interfaces (GUIs) to facilitate the understanding of asymptotic Bode plots and system identification from such plots. The study highlights positive

feedback from users, showcasing how these tools have significantly improved their conceptual understanding and practical application of frequency response techniques. Ramirez-Cortes et al. [4] developed a MATLAB-based educational platform for teaching compensator design in control theory courses. The platform features a graphical user interface with tools for experimenting with controller parameters and visualizing results from both simulated and real signals. Abdulameer et al. [5] emphasized the importance of PID control in process control education. They developed a MATLAB-based GUI software to enhance the teaching and learning of PID control systems. The software allows users to analyze the three-term parameters of a PID controller, explore different control types and structures, and apply tuning methods like Ziegler-Nichols and manual tuning in both the s-domain and z-domain. Valle et al. [6] developed a graphical interface for monitoring key variables in electric vehicles using MATLAB's App Designer environment. Gupta et al. [7] introduced an application, "Analog Parameters Calculator" designed using MATLAB's App Designer to calculate fundamental equations in basic electronics and electrical circuits. The application integrates visual components with programming app behavior, allowing users to compute various electrical parameters, such as Ohm's law, voltage and current division rules, resistor color coding, and circuit analysis for RC, RL, and RLC circuits. Assefa [8] developed a MATLAB App based application to compare the performance of two control techniques, Linear Quadratic Regulator (LQR) and Pole Placement, in DC motor speed control. Rossiter [9] emphasized the importance of active learning in control engineering education by introducing MATLAB-based apps that help students visualize and interact with basic models and system dynamics.

In control systems engineering, transfer functions are fundamental tools used to represent the dynamic behavior of systems. Analyzing these transfer functions often involves decomposing their numerator and denominator polynomials. This decomposition is crucial for gaining a deeper understanding of the transfer functions and for conducting more thorough analyses of system characteristics. Specifically, separating the odd and even parts of the transfer functions is a common method applied in stability analysis, frequency response, and time-domain response studies. Therefore, control systems engineers frequently need to perform such decompositions. For instance, when determining the regions of acceptable PID parameters that ensure system stability, it is necessary to decompose the transfer function into its odd and even components. Many researchers working in control systems utilize the stability boundary locus (SBL) method for this purpose [10]. The SBL method requires the separation of the odd and even parts of both the numerator and denominator of the system's transfer function. In some cases, using these odd and even parts, the real and imaginary parts of the numerator and denominator polynomials are obtained [11]

. As a result, in complex, high-order systems, systems using time delay approximation methods, and systems requiring the conversion of fractional-order systems to integer-order systems, the computational load increases significantly. This, in turn, raises the computational burden on the researcher and complicates the analysis process.

The motivation behind this work was to design a user-friendly interface that, despite the complexity of the plant, allows for the easy decomposition of the numerator and denominator polynomials of a transfer function into their odd and even components. In this study, a MATLAB App has been developed that enables the decomposition of the numerator and denominator polynomials of transfer functions into odd, even, real and imaginary parts. We believe that such a MATLAB App design will enhance the efficiency of researchers by improving the accuracy and speed of the analysis performed on transfer functions.

II. DECOMPOSITION OF THE TRANSFER FUNCTION INTO ODD AND EVEN PARTS

Let the transfer function $T(s)$ be as given in (1), with time delay $\tau \geq 0$.

$$T(s) = \frac{N(s)}{D(s)} e^{-\tau s} = \frac{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0} e^{-\tau s} \quad (1)$$

Where, $N(s)$ and $D(s)$ stand for the numerator and denominator of the transfer function, respectively. In addition, where n and m stand for the order of $N(s)$ and $D(s)$, respectively, with $n \leq m$, b_m and a_n represent the leading coefficients. The numerator and denominator polynomials of a given transfer function can be decomposed into their odd and even components by substituting $s = j\omega$. This process particularly facilitates the use of the SBL method. Additionally, depending on the designer's objectives, this approach can be applied for other purposes as well. After the necessary manipulations, (1) can be written as (2).

$$T(j\omega) = \frac{Ne(-\omega^2) + j\omega No(-\omega^2)}{De(-\omega^2) + j\omega Do(-\omega^2)} e^{-j\omega\tau} \quad (2)$$

Where, Ne and No represent the even and odd components of the numerator polynomial of the transfer function, respectively, while De and Do denote the even and odd components of the denominator polynomial of the transfer function, respectively. Here, it can be observed that the numerator and denominator polynomials, separated into their odd and even components, exhibit a complex structure. Consequently, these polynomials can also be represented in terms of their real and imaginary components. After the necessary transformations, (1) and (2) can be expressed as shown in (3).

$$T(j\omega) = \frac{N_R + jN_I}{D_R + jD_I} e^{-j\omega\tau} \quad (3)$$

Where, $N_R = Ne(-\omega^2)$, $N_I = \omega No(-\omega^2)$, $D_R = De(-\omega^2)$ and $D_I = \omega Do(-\omega^2)$.

III. DESIGNED MATLAB APP FOR DECOMPOSING THE EVEN AND ODD PARTS OF THE TRANSFER FUNCTION

MATLAB App Designer was officially introduced by MathWorks in MATLAB R2016a as a replacement for GUIDE. This modern platform facilitates the creation of graphical user interfaces by providing a streamlined environment for designing and programming applications. App Designer simplifies the interface design process by offering a rich set of standard components, such as gauges, knobs, switches, and lamps, enabling the creation of control panels and human-machine interfaces.

It integrates two key aspects of app development: visually laying out components and programming their functionality, making it an efficient tool for building professional applications. The main interface of App Designer consists of four primary sections: the component library, the design view, the component browser, and the component properties panel [12]. Users can drag and drop elements from the component library to design their interface and use the integrated code editor to define behaviors quickly. The design view allows for interface layout creation, while the code view facilitates functionality programming. The component browser provides centralized management of all used components, and the properties panel enables intuitive adjustment of component settings.

App Designer aligns with modern web design trends, allowing users to develop intuitive, visually appealing, and user-friendly applications. By combining a seamless drag-and-drop interface with powerful coding tools, it offers an all-in-one solution for creating sophisticated GUIs while reducing development time and complexity.

Using the advanced features of this platform, the proposed MATLAB App program was developed using MATLAB R2022a, as outlined below. Fig. 1 shows the main screen of the proposed MATLAB App.

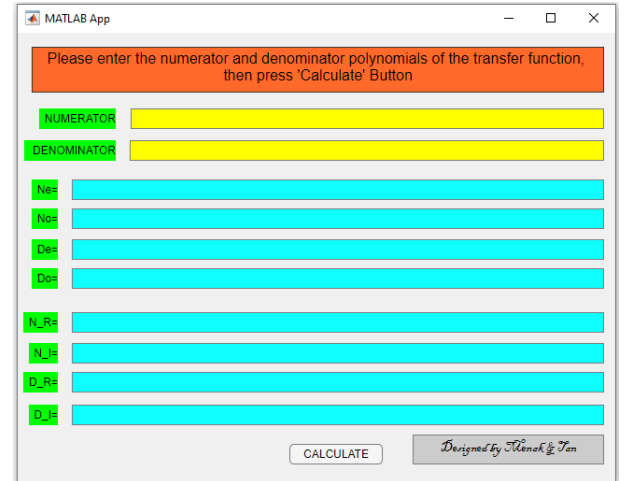


Fig. 1: The main interface of the proposed MATLAB App, designed for decomposing the even, odd, real, and imaginary parts of the numerator and denominator polynomials of a transfer function

The proposed MATLAB App is designed in a simple structure consisting of two user input sections, 8 output sections and one calculation button. The descriptions of the sections are given as follows:

- **NUMERATOR:** This is an edit field text section. This section is designed for the user to input the numerator polynomial of the transfer function.
- **DENOMINATOR:** This is an edit field text section. This section is designed for the user to input the denominator polynomial of the transfer function.
- **Ne:** This is an edit field text section. This field gives the even part of the numerator polynomial.
- **No:** This is an edit field text section. This field gives the odd part of the numerator polynomial.
- **De:** This is an edit field text section. This field gives the even part of the denominator polynomial.
- **Do:** This is an edit field text section. This field gives the odd part of the denominator polynomial.
- **N_R:** This is an edit field text section. This field gives the reel part of the numerator polynomial.
- **N_I:** This is an edit field text section. This field gives the imaginary part of the numerator polynomial.
- **D_R:** This is an edit field text section. This field gives the reel part of the denominator polynomial.
- **D_I:** This is an edit field text section. This field gives the imaginary part of the denominator polynomial.
- **CALCULATE:** This button is used for calculation.

The valid input types for numerator and denominator are expected to be for example as $2*s^2+3*s+1$, $[2\ 3\ 1]$ or $[2, 3, 1]$. Otherwise, the App will display an invalid entry warning, as illustrated in Fig. 2.

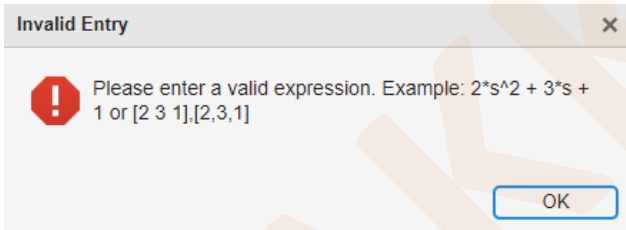


Fig. 2: The pop-up window for the invalid expressions.

IV. IMPLEMENTATION OF THE APP AND EXAMPLES

This section demonstrates how the MATLAB App works and presents the results through the examples given.

A. Example 1

Consider the transfer function given as (4). This is a first order plus time delay system that commonly used in control theory.

$$T(s) = \frac{2}{3s+1} e^{-1s} \quad (4)$$

The numerator and denominator polynomials can be defined as follows:

Numerator : 2 or [2]

Denominator: $3*s+1$ or [3 1]

The implementation and result of the Example 1 are given in Fig. 3.

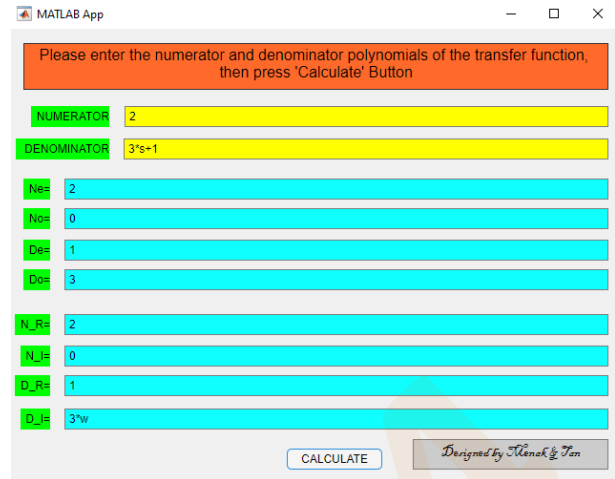


Fig. 3: Implementation and result of the Example 1 via MATLAB App.

While manually calculating a first-order system may be relatively simple and not require the use of the App, it becomes highly beneficial in more complex scenarios. The following examples will be used to illustrate these situations.

B. Example 2

Let consider the high order complicated system [10] as given in (5).

$$T(s) = \frac{N(s)}{D(s)} = \frac{s^3 + 4s^2 - s + 1}{s^5 + 2s^4 + 32s^3 + 14s^2 - 4s + 50} \quad (5)$$

The numerator and denominator polynomials of the Example 2 can be defined as follows:

Numerator : $s^3 + 4*s^2 - s + 1$ or [1 4 -1 1]

Denominator: $s^5 + 2*s^4 + 32*s^3 + 14*s^2 - 4*s + 50$
or [1 2 32 14 -4 50]

The implementation and result of the Example 2 are given in Fig. 4.

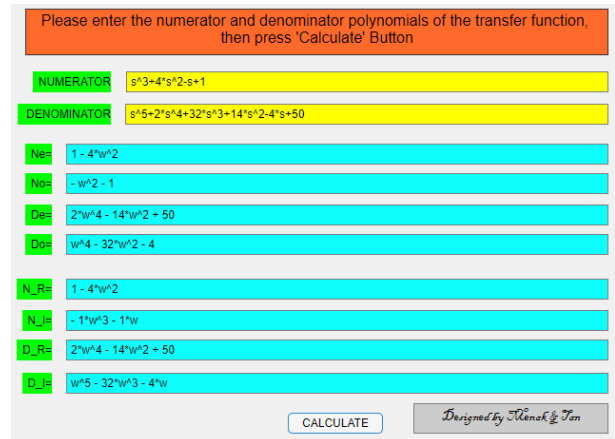


Fig. 4: Implementation and result of the Example 2 via MATLAB App.

C. Example 3

In this example, a high-order time-delay system is considered. The time delay part of the plant can be transformed into numerator and denominator polynomials by some approximation methods. The most commonly used

approximation method is the Padé approximant. Consider the transfer function given in (6).

$$T(s) = \frac{s^2 + 3s + 1}{(s^2 + 20s + 2)(s + 1)^3} e^{-2s} \quad (6)$$

When a third-order Padé approximation is applied [13], the resulting transfer function takes the form shown in (7).

$$T(s) = \frac{\begin{pmatrix} -1s^5 + 3s^4 + 2s^3 - 24s^2 \\ +30s + 15 \end{pmatrix}}{\begin{pmatrix} s^8 + 29s^7 + 218s^6 + 817s^5 + 1748s^4 + \\ 2138s^3 + 1407s^2 + 420s + 30 \end{pmatrix}} \quad (7)$$

The numerator and denominator polynomials of the Example 3 can be defined as follows:

Numerator : [-1 3 2 -24 30 15]

Denominator: [1 29 218 817 1748 2138 1407 420 30]

Using the MATLAB App, the odd, even, real, and imaginary components of the numerator and denominator polynomials of this transfer function are obtained as depicted in Fig. 5.

Please enter the numerator and denominator polynomials of the transfer function, then press 'Calculate' Button

NUMERATOR [-1 3 2 -24 30 15]

DENOMINATOR [1 29 218 817 1748 2138 1407 420 30]

N_e 3*w⁴ + 24*w² + 15

N_o -w⁴ - 2*w² + 30

D_e w⁸ - 218*w⁶ + 1748*w⁴ - 1407*w² + 30

D_o -29*w⁶ + 817*w⁴ - 2138*w² + 420

N_{Re} 3*w⁴ + 24*w² + 15

N_{Im} -1*w⁵ - 2*w³ + 30*w

D_{Re} w⁸ - 218*w⁶ + 1748*w⁴ - 1407*w² + 30

D_{Im} -29*w⁷ + 817*w⁵ - 2138*w³ + 420*w

CALCULATE Designed by *Urenok & Ten*

Fig. 5: Implementation and result of the Example 3 via MATLAB App

D. Example 4

In control systems, fractional-order transfer functions are used alongside integer-order transfer functions. However, during the analysis of fractional-order transfer functions, the fractional-order terms are often converted into integer-order systems using specific transformation techniques. The higher the transformation order, the more accurate the approximation becomes. However, this also significantly increases the order of the transfer function. For such complex systems, manually decomposing the transfer function into its even and odd components becomes highly challenging and time-consuming. With the proposed MATLAB App, this process can be performed effortlessly with a single click. Consider the fractional order transfer function [14], [15] given in (8).

$$T(s) = \frac{1}{s^{1.9} + s} e^{-s} = \frac{1}{s(s^{0.9} + 1)} e^{-s} \quad (8)$$

Consider the 10th-order approximation given in (9) obtained using the Continued Fraction Expansion (CFE) method, as described in [15].

$$s^{0.9} \approx \frac{\begin{pmatrix} 680.37s^{10} + 32586.49s^9 + \\ 409578.54s^8 + \\ 1988381.33s^7 + 4331830.76s^6 + \\ 4493356.65s^5 + \\ 2224971.29s^4 + 498908.57s^3 + \\ 44145s^2 + 1090s + 1 \end{pmatrix}}{\begin{pmatrix} s^{10} + 1090s^9 + 44145s^8 + \\ 498908.57s^7 + 2224971.29s^6 + \\ 4493356.65s^5 + \\ 4331830.76s^4 + 1988381.33s^3 + \\ 409578.54s^2 + 32586.49s + 680.37 \end{pmatrix}} \quad (9)$$

If the term $s^{0.9}$ in (8) is replaced with (9), the numerator and denominator polynomials of the overall system can be expressed as follows:

Numerator: [1 1090.0 44145.0 498908.57 2224971.29
4493356.65 4331830.759 1988381.33
409578.539 32586.49 680.37]

Denominator: [681.373 3676.49 453723.539 20387289.899
6556802.049 8986713.3 6556802.049
2487289.9 453723.539 33676.49 681.37]

Using the MATLAB App, the odd, even, real, and imaginary components of the numerator and denominator polynomials of this transfer function are obtained as depicted in Fig. 6. As can be seen from [15], it is observed that the same results are obtained as expected.

Please enter the numerator and denominator polynomials of the transfer function, then press 'Calculate' Button

NUMERATOR [1 1090.0 44145.0 498908.57 2224971.29 4493356.65 4331830.76 1988381.33 409578.539 32586.49 680.37]

DENOMINATOR [681.373 3676.49 453723.539 20387289.899 6556802.049 8986713.3 6556802.049 2487289.9 453723.539 33676.49 681.37]

N_e -1*w¹⁰ + 44145*w⁸ - 2.2250e+06*w⁶ + 4.3318e+06*w⁴ - 4.0958e+05*w² + 680.3700

N_o 1090*w⁹ - 4.9891e+05*w⁷ + 4.4934e+06*w⁵ - 1.9884e+06*w³ + 3.2586e+04

D_e -3.3676e+04*w¹⁰ + 2.0387e+07*w⁸ - 8.9867e+06*w⁶ + 2.4873e+06*w⁴ - 3.3676e+04*w²

D_o -681.3700*w¹⁰ + 4.5372e+05*w⁸ - 6.5568e+06*w⁶ + 6.5568e+06*w⁴ - 4.5372e+05*w² + 681.3700

N_{Re} -1*w¹⁰ + 44145*w⁸ - 2.2250e+06*w⁶ + 4.3318e+06*w⁴ - 4.0958e+05*w² + 680.3700

N_{Im} 1090*w⁹ - 4.9891e+05*w⁷ + 4.4934e+06*w⁵ - 1.9884e+06*w³ + 3.2586e+04*w

D_{Re} -3.3676e+04*w¹⁰ + 2.0387e+07*w⁸ - 8.9867e+06*w⁶ + 2.4873e+06*w⁴ - 3.3676e+04*w²

D_{Im} -681.3700*w¹¹ + 4.5372e+05*w⁹ - 6.5568e+06*w⁷ + 6.5568e+06*w⁵ - 4.5372e+05*w³ + 681.3700

CALCULATE Designed by *Urenok & Ten*

Fig. 6: Implementation and result of the Example 4 via MATLAB App

V. CONCLUSION

This study presents a MATLAB application designed to facilitate the decomposition of the numerator and denominator polynomials of transfer functions into their even, odd, real, and imaginary components in a systematic manner. This tool enhances the efficiency and accuracy of system analysis by enabling faster and error-free decomposition. The application features a user-friendly interface, significantly improving the efficiency of researchers' analysis of transfer functions. Users can automatically separate the even and odd components of the

polynomials, enabling a more detailed examination of the system's dynamics. This software provides considerable time savings, particularly in the analysis of complex systems, while minimizing potential computational errors. Additionally, the developed application contributes to increasing the efficiency of research in the field of control engineering by enhancing both the speed and accuracy of analytical processes.

Looking ahead, several potential areas for further development of this application exist. Expanding the application to cover broader system models could allow researchers to conduct more complex analyses. Future enhancements could focus on making the user interface more customizable, adding new features to address diverse user needs, and optimizing the software's performance. Furthermore, integrating this application into system modeling and control design processes could transform it into a more comprehensive tool. These improvements would not only enhance the accuracy of research but also provide significant benefits in engineering education and industrial applications.

APPENDIX

In this appendix, we present the function code utilized by the MATLAB App to decompose the numerator and denominator polynomials of a transfer function into their even, odd, real, and imaginary components. Researchers who wish to execute the code as a MATLAB .m file and view the results in the command window can use this code. Please make sure that the function code and the main code are in the same folder.

The function code:

```
function [GNo, GNe, GDo, GDe, GN_R, GN_I, GD_R, GD_I] =
compute_G_odd_even_parts(G_N_coeffs, G_D_coeffs)
    syms s w;

    % Operations for G_N (numerator part)
    % Create symbolic expression from polynomial
    coefficients
    G_N = poly2sym(G_N_coeffs, s);
    coeffs_G_N = coeffs(G_N, s, 'All');
    degree_G_N = length(coeffs_G_N) - 1;

    % Separate odd and even terms
    G_N_odd = 0;
    G_N_even = 0;
    for i = 0:degree_G_N
        term = coeffs_G_N(end - i) * s^i;
        if mod(i, 2) == 0
            G_N_even = G_N_even + term;
        else
            G_N_odd = G_N_odd + term;
        end
    end

    % Substitute s = jw
    GNo = expand(simplify(vpa(subs(G_N_odd, s, j*w)) /
(j*w)));
    GNe = expand(simplify(vpa(subs(G_N_even, s, j*w))));
    GN_R = expand(simplify(vpa(GNe)));
    GN_I = expand(simplify(vpa(w*GNo)));

    % Operations for G_D (denominator part)
    % Create symbolic expression from polynomial
    coefficients
    G_D = poly2sym(G_D_coeffs, s);
    coeffs_G_D = coeffs(G_D, s, 'All');
```

```
degree_G_D = length(coeffs_G_D) - 1;

% Separate odd and even terms
G_D_odd = 0;
G_D_even = 0;
for i = 0:degree_G_D
    term = coeffs_G_D(end - i) * s^i;
    if mod(i, 2) == 0
        G_D_even = G_D_even + term;
    else
        G_D_odd = G_D_odd + term;
    end
end

% Substitute s = jw
GDo = expand(simplify(vpa(subs(G_D_odd, s, j*w)) /
(j*w)));
GDe = expand(simplify(vpa(subs(G_D_even, s, j*w))));
GD_R = expand(simplify(vpa(GDe)));
GD_I = expand(simplify(vpa(w*GDo)));
end
```

This is an example for running the function code above.
The main code:

```
syms s w;
% Polynomial coefficients for the numerator and
denominator
% The numerator polynomial G_N is represented by the
coefficients of the form s^2 + 2*s + 3
G_N = [1 2 3]; % s^2 + 2*s + 3 or [1,2,3]
% The denominator polynomial G_D is represented by the
coefficients of the form s^3 + 2*s + 3
G_D = [1 0 2 3]; % s^3 + 2*s + 3 or [1,0,2,3]
%
% The function will return:
% - GNo: Odd part of the numerator
% - GNe: Even part of the numerator
% - GDo: Odd part of the denominator
% - GDe: Even part of the denominator
% - GN_R: Real part of numerator
% - GN_I: Imaginary part of the numerator
% - GD_R: Real part of the denominator
% - GD_I: Imaginary part of the denominator

[GNo, GNe, GDo, GDe, GN_R, GN_I, GD_R, GD_I] =
compute_G_odd_even_parts(G_N, G_D);

% Display the results, which include the symbolic
expressions for the even and odd parts, as well as the
% real and imaginary components of the transfer
function's numerator and denominator.

disp('GNe:'); % Display the even part of the numerator
disp(GNe);
disp('GNo:'); % Display the odd part of the numerator
disp(GNo);
disp('GDe:'); % Display the even part of the denominator
disp(GDe);
disp('GDo:'); % Display the odd part of the denominator
disp(GDo);
disp('GN_R:'); % Display the real part of the numerator
disp(GN_R);
disp('GN_I:'); % Display the imaginary part of the
numerator
disp(GN_I);
disp('GD_R:'); % Display the real part of the
denominator
disp(GD_R);
disp('GD_I:'); % Display the imaginary part of the
denominator
disp(GD_I);
```

REFERENCES

- [1] Mathworks, "Matlab App Building," 2024. https://www.mathworks.com/help/pdf_doc/matlab/creating_guis.pdf
- [2] A. S. Andreatos and A. D. Zagorianos, "MATLAB GUI application for teaching control systems," in Proceedings of the 6th WSEAS International Conference on ENGINEERING EDUCATION, 2009, no. February 2015, pp. 208–211.
- [3] R. J. Mitchell, "Improved MATLAB GUIs for Learning Frequency Response Methods," IFAC Proc. Vol., vol. 46, no. 17, pp. 303–308, 2013, doi: 10.3182/20130828-3-UK-2039.00067.
- [4] J. M. Ramirez-Cortes, V. Alarcon-Aquino, P. Gomez-Gil, A. Diaz-Mendez, M. Ibarra-Bonilla, and I. Garcia-Enriquez, "Interactive educational tool for compensators design in MATLAB® using frequency response analysis," Comput. Appl. Eng. Educ., vol. 22, no. 4, pp. 699–707, Dec. 2014, doi: 10.1002/cae.21562.
- [5] A. Abdulameer, M. Sulaiman, M. S. M. Aras, and D. Saleem, "GUI Based Control System Analysis using PID Controller for Education," Indones. J. Electr. Eng. Comput. Sci., vol. 3, no. 1, p. 91, Jul. 2016, doi: 10.11591/ijeecs.v3.i1.pp91-101.
- [6] J. M. G. Valle, J. C. C. Garcia, and E. R. Cadaval, "Electric vehicle monitoring system by using MATLAB/App Designer," in 2017 International Young Engineers Forum (YEF-ECE), May 2017, pp. 65–68. doi: 10.1109/YEF-ECE.2017.7935642.
- [7] A. Gupta, P. Kumawat, A. I. Khan, A. Raj, and G. Kaur, "Analog Parameters Calculator using MATLAB App Designer," in IEEE 1st International Conference on "Signal processing, VLSI and Communication Engineering," 2019.
- [8] A. Assefa, "Design GUI App on MATLAB for Comparison Analysis of LQR and Pole Placement Controller for Speed Control DC Motor," Autom. Control Intell. Syst., vol. 11, no. 3, pp. 45–56, Oct. 2023, doi: 10.11648/j.acis.20231103.11.
- [9] J. A. Rossiter, "MATLAB apps to support the learning and understanding of simple system dynamics," IFAC-PapersOnLine, vol. 55, no. 17, pp. 121–125, 2022, doi: 10.1016/j.ifacol.2022.09.267.
- [10] N. Tan, I. Kaya, C. Yeroglu, and D. P. Atherton, "Computation of stabilizing PI and PID controllers using the stability boundary locus," Energy Convers. Manag., vol. 47, no. 18–19, pp. 3045–3058, Nov. 2006, doi: 10.1016/j.enconman.2006.03.022.
- [11] R. Menak and N. Tan, "Computation of All Robustly Stabilizing PID Controllers Based on H_∞ Robust Stability Condition," IEEE Access, vol. 12, no. October, pp. 140556–140593, 2024, doi: 10.1109/ACCESS.2024.3461315.
- [12] G. Ren et al., "Simulation Analysis for Electromagnetic Emission Characteristics of Electric Vehicles Based on MATLAB App Designer," in 2022 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC), Sep. 2022, pp. 258–260. doi: 10.1109/APEMC53576.2022.9888700.
- [13] M. Vajta, "Some remarks on Padé -approximations," in Proceedings of the 3rd TEMPUS-INTCOM Symposium, 2000, vol. 242, pp. 1–6.
- [14] M. M. Ozyetkin and N. Tan, "Practical Tuning Algorithm of PD μ Controller for Processes with Time Delay," IFAC-PapersOnLine, vol. 50, no. 1, pp. 9230–9235, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.1281.
- [15] M. M. Özyetkin and D. Astekin, "Fractional Order PD Controller Design Using Substitutions of Time Delay," in Theory and Research in Engineering, 2022, pp. 57–72.

Endüstriyel Uygulamalarda Blok Zincir Türlerinin Kullanımı: Verimlilik, Güvenlik ve Ölçeklenebilirlik Değerlendirmesi

The Use of Blockchain Types in Industrial Applications: An Evaluation of Efficiency, Security, and Scalability

1th Burak Ağgöl

Bilişim Teknolojileri Bölümü
Sakarya Uygulamalı Bilimler
Üniversitesi

Sakarya, Türkiye

burakaggul@subu.edu.tr

Bilgisayar Bilimleri ve Mühendisliği
İstanbul Sabahattin Zaim Üniversitesi
İstanbul, Türkiye

aggul.burak@std.izu.edu.tr

2th Tayfun Acarer

Motorlu Araçlar ve Ulaştırma
Teknolojileri Bölümü
Piri Reis Üniversitesi

İstanbul, Türkiye

tacarar@pirireis.edu.tr

3th Gökhan Erdemir

Department of Engineering
Management and Technology
The University of Tennessee at
Chattanooga

Chattanooga, TN, USA

gokhan-erdemir@utc.edu

Özetçe— Blok zincir teknolojisi, verimlilik, güvenlik ve izlenebilirlik sağlayarak maliyetleri düşürmekte ve endüstriyel otomasyon gibi alanlarda çözümler sunmaktadır. Diğer taraftan, blok zincirler farklı altyapıları nedeniyle her uygulama için uygun olmayabilir. Endüstriyel otomasyon uygulamalarında enerji gereksinimi, güvenlik, ölçeklenebilirlik ve performans gibi faktörler dikkate alınarak doğru blok zincir seçimi yapılmalıdır. Bu çalışma, blok zincir türlerinin endüstriyel uygulamalardaki uygunluğunu ve performansını değerlendirerek araştırmacılar ve geliştiriciler için rehber niteliğindedir.

Anahtar Kelimeler — blokzincir; endüstriyel uygulamalar; verimlilik; güvenlik.

Abstract—Blockchain technology reduces costs by providing efficiency, security, and traceability and offers solutions in areas such as industrial automation. On the other hand, blockchains may not be suitable for every application due to their different infrastructures. In industrial automation applications, the proper blockchain selection should be made by considering factors such as energy requirements, security, scalability, and performance. This study evaluates the suitability and performance of blockchain types in industrial applications and is a guide for researchers and developers.

Keywords — blockchain; industrial applications; efficiency; security.

I. GİRİŞ

Blok zinciri teknolojisinin ilk kullanımı, blok zinciri teknolojisine dayanan ve çevrimiçi olarak mal alışverişi yapmak için geleneksel para değişimlerine benzer bir yöntem kullanan dijital bir para birimi olan Bitcoin'dir. Bitcoin'in popülaritesi sayesinde, insanlar artık finansal piyasa, nesnelerin interneti (IoT), tedarik zinciri, oylama, tıbbi tedavi ve depolama gibi geniş bir alan ve hizmet yelpazesinde blok zinciri teknolojilerini kullanabilirler. 2008'de blok zincirinin ilk merkeziyetsiz kripto para biriminin temeli olarak

tanıtılması, güvenli, etkili ve şeffaf bir eşler arası bilgi alışverişine olanak sağladı.

Blok zinciri, işlemleri kronolojik sırayla kaydeden ve uygun bir mutabakat yöntemiyle güvence altına alınan, herkese açık, değiştirilemez bir defterdir. Öne çıkan özellikleri arasında değişmezlik, geri döndürülemezlik, merkeziyetsizlik, kalıcılık ve anonimlik yer alır. Blok zincir, kronolojik olarak birbirine bağlı ve fikir birliği mekanizmalarıyla güvence altına alınmış işlem verilerini içeren bloklardan oluşur [1]. Blok zinciri, verileri dağıtık bir şekilde depolayarak sürekli artan kayıtları koruma altına alan bir sistemdir. Programları ve işlemleri içeren bloklardan oluşur ve her işlem ağı tamamında görülebilir. Her düğümün tüm blok zincir'in bir kopyasını içerdiği, veri bütünlüğünü ve değiştirilemezliğini garanti eden eşler arası bir ağ üzerinde çalışır [2].

Blok zinciri, iş ve endüstri dünyası için yeni ve gelişmiş özellikler sunarken aynı zamanda çeşitli endüstrilerdeki sorunları çözmeye yetenekleri nedeniyle son zamanlarda büyük ilgi görmektedir. Bu teknolojiler, güvenli veri kaydı ve paylaşımı, tedarik zinciri süreçlerinin otomasyonu ve şeffaflığın artırılması gibi sorunları ele almak için etkili bir yol sunar. Blok zincirinin kullanımı, verimlilik ve güvenliği artırırken, izlenebilirlik ve şeffaflığı geliştirir ve maliyetleri düşürerek birçok endüstriyel uygulamayı mümkün kılar [3]. Blok zinciri teknolojisi, eşler arası ağların ve kriptografik algoritmaların avantajlarını birleştirerek anlaşmaların geçerliliğini garanti eder. Tüm katılımcı varlıklar, diğer katılımcı varlıkları dahil etmeden onaylanmış ve kaydedilmiş bir faaliyeti değiştiremez. Sonuç olarak, çeşitli sektörlerde ve işletmelerde blok zinciri kullanımını değerlendirmek ve bunların bu sektörlerde etkili bir şekilde uygulanabilmesi için daha fazla araştırma yapılmaktadır [3].

Bir blok zincir platformu seçerken, kuruluşlar işlevsellik, uyarlanabilirlik ve mevcut sistemlerle uyumluluk gibi faktörleri göz önünde bulundurmalıdır [4]. Blok zincir mimarilerinin ve fikir birliği protokollerinin evrimi, her biri farklı uygulamalara uygun benzersiz özelliklere sahip bir dizi

platforma yol açmıştır [5]. Tedarik zinciri şeffaflığı ve sürdürülebilirliği için, blok zincir teknolojisi değerlendirmesi hem teknik hem de sürdürülebilir nitelikleri içermeli ve karar alma süreçleri psikolojik faktörleri ve farklı görüşleri hesaba katmalıdır [6]. Her birinin güçlü ve zayıf yönleri olan blok zincir türlerinin çeşitliliği, en uygun çözümü seçerken proje gereksinimlerinin dikkatlice değerlendirilmesini gerektirir [5], [4]. Bu nedenle, tek doğru seçim meselesi kullanım senaryosuna bağlıdır. Son olarak blok zinciri hakkında birkaç iyi inceleme makalesi bulunmasına rağmen, bu makaledeki kapsamımız ve amaçlarımız farklıdır; blok zincirinin endüstriyel uygulamalarda etkili bir şekilde kullanılmasına yönelik blok zincir türleri ve teknolojilerini kıyaslıyoruz.

Bu çalışma, blok zincir türlerinin enerji tüketimi, güvenlik, ölçeklenebilirlik ve performans açısından karşılaştırmalı analizini sunarak, araştırmacılara ve geliştiricilere yol göstermeyi amaçlamaktadır. Özellikle endüstriyel otomasyon uygulamalarında, hangi blok zincir platformlarının ve algoritmalarının daha uygun olduğuna karar verirken dikkate alınması gereken çeşitli faktörler bulunmaktadır. Bu doğrultuda, çalışmada aşağıdaki sorulara cevap aranmaktadır:

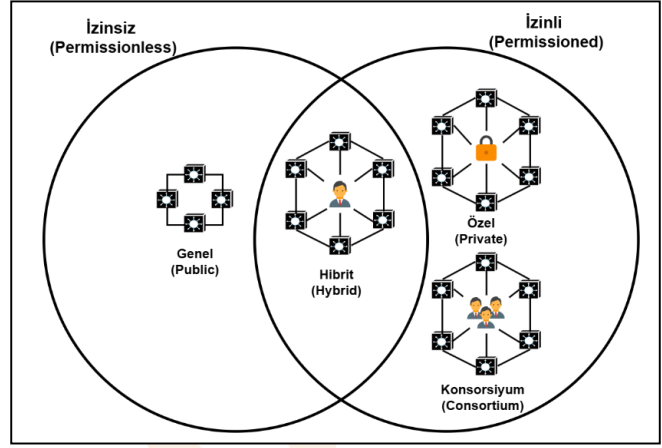
- Blok zincir türleri (açık, kapalı, hibrit ve konsorsiyum) arasında performans, güvenlik ve ölçeklenebilirlik açısından ne gibi farklar vardır?
- Hangi blok zincir türleri, endüstriyel otomasyon uygulamaları için en uygun platformları sunmaktadır?
- Blok zincir teknolojilerinde kullanılan Mutabakat algoritmalarının enerji verimliliği ve işlem hızı açısından nasıl bir performans göstermektedir?
- Açık blok zincirler ile kapalı blok zincirler arasında hangi kullanım alanları için önemli farklar bulunmaktadır?
- Hibrit ve konsorsiyum blok zincirlerinin avantajları nelerdir ve hangi uygulamalarda tercih edilirler?

Bu sorular çerçevesinde, blok zincir türlerinin farklı endüstriyel uygulamalardaki uygunluğu ve performansı değerlendirilecektir. Çalışmanın amacı, blok zincir teknolojisinin hangi türünün ve platformunun belirli kullanım senaryoları için daha avantajlı olduğunu belirlemektir.

II. BLOK ZİNCİR TÜRLERİ

Literatürü incelediğimizde temelde iki tür blokzincir bulunmaktadır diyen araştırmacılar tarafından, izinsiz (permission-less) ve izinli (permissioned)[7], [8], [9]. Bu iki tür blok zincirin, farklı çalışmalarda Kamu (Public) ve Özel (Private) blok zincirleri olarak adlandırıldığı belirtilmiştir[10], [9]. Aslında iki ayrı adlandırılma olarak gözüke de birbirleri yerine kullanılan ifadelerin aslında aynı tür blok zincirleri temsil ettiğini görmekteyiz [11], [12], [9]. Gizlilik, güvenlik gibi gereksinimlerin karşılanabilmesi için bu iki model, blok zinciri dünyasında büyük önem taşımaktadır [13]. Fakat Attaran ve arkadaşlarına göre dört tür blok zincir bulunmaktadır [14]. Daha detaylı literatür taraması yapıldığında bu farklı fikirlerin sebebinin aslında başlangıçta, blok zinciri modelleri izinsiz (permissionless) ve

izinli (permissioned) olmak üzere iki temel kategoriye ayrılırken, teknolojik ilerlemeler ve sektörel ihtiyaçlar doğrultusunda bu sınıflandırma hibrit ve konsorsiyum blok zincirleri gibi üç ya da dört farklı türe genişlediği bilgisi anlaşılmıştır [15-19]. Bu genişlemenin nedeni, teknoloji geliştikçe kullanım alanlarının, teknik detayların ve sistem gereksinimlerinin çeşitlenmesidir [20], [21]. Bu farklılığının sınıflandırıldığı sonucu Şekil 1’de görebilirsiniz.



Şekil 1. Blok zincir Türleri ve Sınıflandırılması [22] [23]

Burada yola çıkılarak çeşitli uygulamaların gereksinimlerini karşılamak amacıyla, blok zincirleri genellikle diğer düğümlerin başlattığı işlemlere hangi düğümlerin erişebileceğini, bu işlemleri doğrulayabileceğini ve kimlik doğrulaması yapabileceğini belirleyen kurallara göre yapılandırılır [15],[24]. Blok zinciri yapılarının karşılaştırması Tablo 1'de verilmiştir. Bu bağlamda ortaya çıkan blok zincir türlerini kısaca açıklamak gerekirse:

A. Açık veya İzinsiz Blok Zincirleri

Açık veya izinsiz blok zincirleri, herkesin katılımına ve ağ üzerinde işlem yapmasına izin veren merkeziyetsiz sistemlerdir [7], [15]. Bu tür blok zincirlerinde, herhangi bir kullanıcı ağa dahil olabilir, işlemleri doğrulayabilir ve yeni bloklar ekleyebilir [10]. Şeffaflık ve güvenlik, Mutabakat algoritmaları aracılığıyla sağlanır. Bitcoin ve Ethereum gibi popüler kripto para birimleri, bu model üzerinde çalışmaktadır [16], [8].

B. Özel veya İzinli Blok Zincirleri

Özel veya izinli blok zincirleri, erişimin ve katılımın belirli bir organizasyon veya grup tarafından kontrol edildiği ağlardır [7], [15], [8].

Bu sistemlerde, yalnızca yetkilendirilmiş kullanıcılar işlemleri görüntüleyebilir ve doğrulayabilir [10]. Yüksek düzeyde gizlilik ve kontrol sağladığı için genellikle kurumsal çözümlerde tercih edilir. Örnek olarak, Hyperledger Fabric ve R3 Corda bu modele uygundur[16], [8].

C. Konsorsiyum Blok Zincirleri

Konsorsiyum blok zincirleri, birden fazla organizasyonun ortaklaşa yönettiği ve kontrol ettiği yarı özel ağlardır [15],[18]. Katılım ve Mutabakat mekanizmaları, önceden belirlenmiş bir grup tarafından yönetilir[18], [16].

Bu yapı, farklı kurumların iş birliği yaparken hem veri paylaşımını hem de gizliliği dengeli bir şekilde

yönetmelerine olanak tanır [19]. Bankacılık ve finans sektöründeki bazı ortak platformlar bu modeli kullanmaktadır.

D. Hibrit Blok Zincirleri

Hibrit blok zincirleri, açık ve özel blok zincirlerinin özelliklerini bir araya getiren karma sistemlerdir [20]. Bu tür ağlarda, bazı veriler ve işlemler herkese açıkken, diğerleri sadece yetkili kullanıcılar tarafından erişilebilir [16].

Bu sayede hem şeffaflık hem de gizlilik ihtiyaçları aynı anda karşılanabilir. Hibrit modeller, çeşitli sektörlerin spesifik gereksinimlerine esnek çözümler sunar.

Tablo 1, blockchain türlerinin erişim yapısı, merkeziyetsizlik seviyesi, veri modeli, güvenlik ve kullanım senaryoları gibi özellikleri kapsar [22].

TABLO I. BLOK ZİNCİR TÜRLERİNİN KARAKTERİSTİK KARŞILAŞTIRILMASI

Blok Zincir Türü	Karakteristik Özellikler				
	Ağ Erişimi	Merkeziyetsizlik	Veri Yapısı	Güvenlik	Kullanım Alanları
Public	İzinsiz	Yüksek	UTXO, Account-Based	Düşük	Kripto Para, Akıllı Sözleşmeler
Consortium	İzinli	Orta	Account-Based	Orta	Bankacılık, Tedarik Zinciri
Hybrid	İzinsiz/İzinli	Orta	UTXO, Account-Based	Yüksek	IoT, Kurumsal Blockchain Çözümleri
Private	İzinli	Düşük	Account-Based	Yüksek	Kurum içi çözümler, Kurumsal

Kaynak:[25-30]

III. MUTABAKAT ALGORİTMALARI

Bir blok zincir ağında, bir işlemin başarılı sayılabilmesi için iki kez doğrulanması gerekir. İki taraf arasındaki işlemi kolaylaştırmak için ilk aşamada bir akıllı sözleşme kullanılır. İşlemin mevcut durumu üzerinde anlaşmaya varmak için ise ikinci aşamada bir Mutabakat algoritması kullanılır. Blok zincir teknolojisinin temeli akıllı sözleşmelerdir. İki taraf bir akıllı sözleşmeyi imzaladığında ve tüm şartlarını kabul ettiğinde, sözleşme anında blok zincir'e yüklenir ve belirli koşullar sağlandığında işlem tamamlanır [31].

Bilgisayar tabanlı işlem protokolleri olarak bilinen akıllı sözleşmeler, yasal olayları kaydetme, yönetme, yürütme ve sözleşme şartlarına uygun olarak faaliyetlerde bulunma yeteneğine sahiptir [32]. Akıllı sözleşmelerin blok zincir teknolojisine dahil edilmesinin başlıca amaçları, dolandırıcılık kayıplarını azaltmak, güvenilir araçların gerekliliğini ortadan kaldırmak ve yürütme ile beklenmeyen istisnalarla ilgili masrafları düşürmektir [33].

İkinci aşamada anlaşma süreci, Mutabakat olarak bilinir; her düğüm, mevcut bir blok zincire yeni bir blok ekleme yöntemini seçer. Blok zincir tabanlı herhangi bir sistemin temeli Mutabakat mekanizmasıdır; bir sistem, onu yöneten Mutabakat protokolü kadar güçlü ve güvenilirdir [34]. Yeni bir blok eklemek için anlaşmaya varmak amacıyla, katılımcı düğümler çeşitli Mutabakat teknikleri kullanabilir [35], [36].

Mutabakat protokolleri, çalışma şekillerine göre geniş çapta gruplanabilir ve protokolleri, oylamaya dayalı veya kanıta dayalı olarak sınıflandırılabilir [37], [38]. Oylamaya

dayalı protokoller izinli blok zincirlerde kullanılırken, kanıta dayalı Mutabakat yöntemleri izinsiz blok zincirlerin temelini oluşturur.

Dağıtık veya merkeziyetsiz ağlarda mutabakata ulaşmak için herhangi bir dizi kriter, örneğin çoğunluk yapısı, merkeziyetsiz yönetim, işlem gücü, Bizans hata toleransı vb. kullanılabilir. Bitcoin gibi kamuya açık blok zincirlerde, "İş Kanıtı" (Proof of Work) fikri benimsenmiştir. Konsorsiyum ve özel blok zincirlerde, oylamaya dayalı Mutabakat protokollerinin (PBFT, RAFT, Ripple ve Stellar gibi) yerine, "Varoluş Kanıtı" (Proof of Existence - PoE), "Hisse Kanıtı" (Proof of Stake - PoS), "Geçen Süre Kanıtı" (Proof of Elapsed Time - PoET), "Yetkilendirilmiş Hisse Kanıtı" (Delegated Proof of Stake - DPoS), "Önem Kanıtı" (Proof of Importance), "Depolama Kanıtı" (Proof of Storage) ve diğer hibrit kanıta dayalı mutabakat algoritmaları tercih edilmektedir [39].

IV. ENDÜSTRİYEL UYGULAMALARDA BLOK ZİNCİR KULLANIMI

Tablo 2' ye göre, açık kaynaklı blok zincir platformları arasında Ethereum ve Hyperledger Sawtooth gibi platformlar, akıllı sözleşmelerin ve merkeziyetsiz uygulamaların geliştirilmesine olanak tanırken, IOTA daha çok IoT odaklı ve enerji verimliliği yüksek bir platform olduğu belirtilmektedir. Kapalı kaynaklı blok zincirler arasında Corda ve Quorum gibi platformlar, yüksek güvenlik ve enerji verimliliği sunarken, genellikle kurumsal ve finansal uygulamalar için tercih edilmektedir. Hibrit blok zincirler hem merkeziyetsizliği hem de güvenliği dengelemeye çalışan platformlar olup, Dragonchain ve XinFin gibi platformlar akıllı sözleşmeler ve kurumsal uygulamalarda yaygın olarak kullanılmaktadır. Konsorsiyum blok zincirleri ise birden fazla kuruluşun iş birliği yaptığı sistemler olup, Hyperledger Fabric ve VeChain gibi platformlar tedarik zinciri ve lojistik gibi sektörlerde güçlü bir izlenebilirlik ve güvenlik sağlamaktadır.

V. BLOK ZİNCİR TÜRLERİNİN ENDÜSTRİYEL PERFORMANS ANALİZİ

Tablo 2' de yer alan veriler, endüstriyel otomasyon sistemlerinde kullanılan dört farklı blok zinciri türüne (açık kaynaklı, kapalı kaynaklı, hibrit ve konsorsiyum) göre en sık kullanılan platformları ve bunların çeşitli özelliklerini karşılaştırmaktadır.

Her platformun mutabakat algoritması, kullanım alanları, işlem hızı (TPS), güvenlik, ölçeklenebilirlik, enerji verimliliği, akıllı sözleşmelerin (Smart Contracts) desteklenmesi, merkeziyetsiz uygulamaların (DApps) geliştirilip geliştirilebilmesi, açık kaynak kodlu olup olmaması ve açık API'ler (Open APIs) sunulup sunulmadığı gibi önemli özellikleri içermektedir [36], [40-49]. Tablo 2, oluşturulurken [28], [50-61] ilgili akademik çalışmalar incelenmiştir.

TABLO II. BLOK ZİNCİR PLATFORMLARI VE KULLANDIKLARI MUTABAKAT ALGORİTMALARININ KARŞILAŞTIRMASI

Blok Zincir Türü	Karşılaştırılan Özellikler										
	Platform	Mutabakat Algoritması	Kullanım Alanları	İşlem Hızı (TPS)	Güvenlik	Ölçeklenebilirlik	Enerji Verimliliği	Smart Contract	DApps	Open Source	Open APIs
Açık Kaynaklı Blok Zincirler	Ethereum	Proof of Stake (PoS)	Akıllı sözleşmeler, finans, tedarik zinciri	~15 TPS	Orta	Orta	Yüksek	Evet	Evet	Evet	Evet
	Hyperledger Sawtooth	Proof of Elapsed Time (PoET)	IoT, tedarik zinciri, endüstriyel otomasyon	~1000 TPS	Yüksek	Yüksek	Orta	Evet	Hayır	Evet	Evet
	IOTA	Tangle (DAG)	IoT, makine ödemeleri	Yüksek (Teorik olarak sınırsız)	Orta	Çok Yüksek	Yüksek	Hayır	Hayır	Evet	Evet
Kapalı Kaynaklı Blok Zincirler	Corda	Noter Tabanlı Mutabakat	Finans, kurumsal otomasyon	~170 TPS	Çok Yüksek	Orta	Yüksek	Hayır	Hayır	Hayır	Hayır
	Quorum	Proof of Authority (PoA), Istanbul BFT	Bankacılık, finansal hizmetler, kurumsal çözümler	~100 TPS	Yüksek	Orta	Yüksek	Evet	Evet	Hayır	Evet
	Ripple	Ripple Protocol Mutabakat Algorithm (RPCA)	Finansal işlemler, bankalar arası para transferi	~1500 TPS	Yüksek	Orta	Yüksek	Hayır	Hayır	Hayır	Evet
Hibrit Blok Zincirler	Polkadot	Nominated Proof of Stake (NPoS)	Akıllı sözleşmeler, IoT, tedarik zinciri	~1000 TPS	Yüksek	Yüksek	Yüksek	Evet	Evet	Evet	Evet
	XinFin (XDC Network)	Delegated Proof of Stake (DPoS)	Ticaret ve finans, IoT	~2000 TPS	Yüksek	Yüksek	Yüksek	Evet	Evet	Evet	Evet
	Kadena	Proof of Work (PoW), Chainweb	Akıllı sözleşmeler, endüstriyel otomasyon	~480,000 TPS (Teorik)	Yüksek	Yüksek	Orta	Evet	Hayır	Evet	Evet
Konsorsiyum Blok Zincirler	R3 Corda	Noter Tabanlı Mutabakat	Finans, bankacılık, otomasyon sistemleri	~170 TPS	Çok Yüksek	Orta	Yüksek	Hayır	Hayır	Hayır	Hayır
	Hyperledger Fabric	Practical Byzantine Fault Tolerance (PBFT)	Tedarik zinciri, kurumsal çözümler	~3500 TPS	Yüksek	Yüksek	Orta	Evet	Hayır	Evet	Evet
	VeChain	Proof of Authority (PoA)	Tedarik zinciri, lojistik, endüstriyel uygulamalar	~10,000 TPS	Yüksek	Yüksek	Yüksek	Evet	Hayır	Hayır	Evet

VI. SONUÇ

Bu çalışma, dört ana blokzincir türünün (açık, kapalı, hibrit ve konsorsiyum) endüstriyel uygulamalardaki kullanımlarını ve performanslarını sayısal verilerle detaylandırarak analiz etmiştir.

Açık blokzincirler, merkeziyetsizlik ve güvenliği ön planda tutan yapılar olup, kripto para ve merkeziyetsiz uygulamalar (DeFi) gibi kullanım alanlarında yaygın olarak tercih edilmektedir. Ancak bu platformlar, özellikle Proof of

Work (PoW) algoritmasına dayandıkları için yüksek enerji tüketimi ile dikkat çekmektedir. Bu platformlar, şeffaflık ve güven gerektiren uygulamalarda avantaj sağlamaktadır, ancak ölçeklenebilirlik ve işlem hızı açısından sınırlamalara sahiptir. Kapalı blokzincirler, güvenlik ve gizlilik gereksinimi olan kurumsal çözümler için idealdir. Bu platformlar, sınırlı erişimle çalıştıkları için enerji verimliliği yüksek olup, güvenlik ve ölçeklenebilirlik açısından avantaj sağlamaktadır. Hibrit blokzincirler, açık ve kapalı blokzincirlerin avantajlarını birleştirir ve hem performans hem de güvenlik sağlar. Bu platformlar, özellikle IoT, ticaret ve finans uygulamalarında kullanılarak hem kamuya açık

işlemlerin şeffaflığını hem de özel işlemlerin gizliliğini sağlamaktadır. Proof of Stake (PoS) ve Proof of Authority (PoA) gibi enerji verimli Mutabakat algoritmaları, hibrit blokzincirlerin düşük enerji tüketimi ve yüksek güvenlik sunmasına olanak tanımaktadır. Konsorsiyum blok zincirleri, birden fazla kuruluşun iş birliği yaptığı yarı kapalı blok zincir yapıları sunmaktadır.

Genel olarak, blok zincir türüne ve kullanım amacına bağlı olarak algoritmaların enerji tüketimi, güvenlik, ölçeklenebilirlik ve performans açısından avantajları ve dezavantajları bulunmaktadır. Bu çalışma, her iki bakış açısından da blokzincir sistemlerinin özelliklerini ele alan iki temel tablo sunmaktadır. İlk tablo, platform seviyesinde bir inceleme sunarken, ikinci tablo algoritmaların teknik detaylarına odaklanarak karşılaştırma yapmaktadır. Platformların farklı kullanım alanları ve teknik özellikleri göz önüne alındığında, enerji verimliliği, işlem hızı ve güvenlik gibi faktörler, belirli bir uygulama için en uygun blok zincir teknolojisinin seçilmesinde kritik bir rol oynar. Endüstriyel otomasyon uygulamalarına yönelik en uygun blok zincir platformlarını anlamak isteyen araştırmacı ve geliştiriciler için bu tür tablolar değerli bir rehber niteliğindedir. Sonuç olarak, platformların performansı, güvenliği ve enerji verimliliği gibi özellikleri dikkate alınarak yapılacak doğru bir seçim, projenin başarısını doğrudan etkileyen önemli bir unsurdur.

KAYNAKLAR

- [1] D. Puthal, N. Malik, S. P. Mohanty, E. Kougiannos, and G. Das, "Everything You Wanted to Know About the Blockchain: Its Promise, Components, Processes, and Problems," *IEEE Consumer Electronics Magazine*, vol. 7, no. 4, pp. 6–14, Jul. 2018, doi: 10.1109/MCE.2018.2816299.
- [2] V. Patel, F. Khatawala, K. Shah, and Y. Choksi, "A Review on Blockchain Technology: Components, Issues and Challenges," 2020, pp. 1257–1262. doi: 10.1007/978-981-15-1420-3_137.
- [3] J. Al-Jaroodi and N. Mohamed, "Blockchain in Industries: A Survey," *IEEE Access*, vol. 7, pp. 36500–36515, 2019, doi: 10.1109/ACCESS.2019.2903554.
- [4] S. Farshidi, S. Jansen, S. Espana, and J. Verkleij, "Decision Support for Blockchain Platform Selection: Three Industry Case Studies," *IEEE Trans Eng Manag*, vol. 67, no. 4, pp. 1109–1128, Nov. 2020, doi: 10.1109/TEM.2019.2956897.
- [5] L. Ismail and H. Materwala, "A Review of Blockchain Architecture and Consensus Protocols: Use Cases, Challenges, and Solutions," *Symmetry (Basel)*, vol. 11, no. 10, p. 1198, Sep. 2019, doi: 10.3390/sym11101198.
- [6] C. Bai and J. Sarkis, "A supply chain transparency and sustainability technology appraisal model for blockchain technology," *Int J Prod Res*, vol. 58, no. 7, pp. 2142–2162, Apr. 2020, doi: 10.1080/00207543.2019.1708989.
- [7] D. Li, W. E. Wong, and J. Guo, "A Survey on Blockchain for Enterprise Using Hyperledger Fabric and Composer," in *Proceedings - 2019 6th International Conference on Dependable Systems and Their Applications, DSA 2019*, Institute of Electrical and Electronics Engineers Inc., Jan. 2020, pp. 71–80. doi: 10.1109/DSA.2019.00017.
- [8] K. Wust and A. Gervais, "Do you need a blockchain?," in *Proceedings - 2018 Crypto Valley Conference on Blockchain Technology, CVCBT 2018*, Institute of Electrical and Electronics Engineers Inc., Nov. 2018, pp. 45–54. doi: 10.1109/CVCBT.2018.00011.
- [9] G. Falazi, M. Hahn, U. Breitenbacher, F. Leymann, and V. Yussupov, "Process-based composition of permissioned and permissionless blockchain smart contracts," in *Proceedings - 2019 IEEE 23rd International Enterprise Distributed Object Computing Conference, EDOC 2019*, Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 77–87. doi: 10.1109/EDOC.2019.00019.
- [10] M. S. Ferdous, M. J. M. Chowdhury, M. A. Hoque, and A. Colman, "Blockchain Consensus Algorithms: A Survey," Jan. 2020, [Online]. Available: <http://arxiv.org/abs/2001.07091>
- [11] S. Solat, P. Calvez, and F. Nait-Abdesslem, "Permissioned vs. Permissionless Blockchain: How and Why There Is Only One Right Choice," *Journal of Software*, pp. 95–106, May 2021, doi: 10.17706/jsw.16.3.95-106.
- [12] C. V. Helliari, L. Crawford, L. Rocca, C. Teodori, and M. Veneziani, "Permissionless and permissioned blockchain diffusion," *Int J Inf Manage*, vol. 54, Oct. 2020, doi: 10.1016/j.ijinfomgt.2020.102136.
- [13] N. Kshetri, "Blockchain's roles in strengthening cybersecurity and protecting privacy," *Telecomm Policy*, vol. 41, no. 10, pp. 1027–1038, Nov. 2017, doi: 10.1016/j.telpol.2017.09.003.
- [14] M. Attaran and A. Gunasekaran, *Applications of Blockchain Technology in Business*. in SpringerBriefs in Operations Management. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-27798-7.
- [15] A. Guru, B. K. Mohanta, H. Mohapatra, F. Al-Turjman, C. Altrjman, and A. Yadav, "A Survey on Consensus Protocols and Attacks on Blockchain Technology," Feb. 01, 2023, *MDPI*. doi: 10.3390/app13042604.
- [16] S. Santhoshi, D. Arigela, and P. Voolla, "Blockchain Open Source Tools: Ethereum and Hyperledger Fabric."
- [17] D. Chatziamanetoglou and K. Rantos, "Cyber Threat Intelligence on Blockchain: A Systematic Literature Review," Mar. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/computers13030060.
- [18] N. Verma, S. Jain, and R. Doriya, "Review on consensus protocols for blockchain," in *Proceedings - IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2021*, Institute of Electrical and Electronics Engineers Inc., Feb. 2021, pp. 281–286. doi: 10.1109/ICCIS51004.2021.9397089.
- [19] W. Yao, F. P. Deek, R. Murimi, and G. Wang, "SoK: A Taxonomy for Critical Analysis of Consensus Mechanisms in Consortium Blockchain," *IEEE Access*, vol. 11, pp. 79572–79587, 2023, doi: 10.1109/ACCESS.2023.3298675.
- [20] H. M. Kim, H. Turesson, M. Laskowski, and A. F. Bahreini, "Permissionless and Permissioned, Technology-Focused and Business Needs-Driven: Understanding the Hybrid Opportunity in Blockchain Through a Case Study of Insolar," *IEEE Trans Eng Manag*, vol. 69, no. 3, pp. 776–791, Jun. 2022, doi: 10.1109/TEM.2020.3003565.
- [21] P. Fraga-Lamas, T. M. Fernández-Caramés, A. M. Rosado Da Cruz, and S. I. Lopes, "An Overview of Blockchain for Industry 5.0: Towards Human-Centric, Sustainable and Resilient Applications," 2024, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2024.3435374.
- [22] B. Guidi and A. Michienzi, "From NFT 1.0 to NFT 2.0: A Review of the Evolution of Non-Fungible Tokens," Jun. 01, 2023, *MDPI*. doi: 10.3390/fi15060189.
- [23] S. Islam, M. J. Islam, M. Hossain, S. Noor, K. S. Kwak, and S. M. R. Islam, "A Survey on Consensus Algorithms in Blockchain-Based Applications: Architecture, Taxonomy, and Operational Issues," *IEEE Access*, vol. 11, pp. 39066–39082, 2023, doi: 10.1109/ACCESS.2023.3267047.
- [24] A. Rahman *et al.*, "On the Integration of Blockchain and SDN: Overview, Applications, and Future Perspectives," *Journal of Network and Systems Management*, vol. 30, no. 4, Oct. 2022, doi: 10.1007/s10922-022-09682-4.
- [25] B. Assiri and H. Alnami, "Smart Partitioned Blockchain," *Sensors*, vol. 24, no. 13, Jul. 2024, doi: 10.3390/s24134033.
- [26] V. Sharma and N. Lal, "A NOVEL COMPARISON OF CONSENSUS ALGORITHMS IN BLOCKCHAIN," 2020.
- [27] *ICOIN 2018 : the 32nd International Conference on Information Networking : January 10-12, 2018, Holiday Inn ChiangMai, Chiang Mai, Thailand*. Institute of Electrical and Electronics Engineers, 2018.
- [28] Y. Merrad *et al.*, "Blockchain: Consensus Algorithm Key Performance Indicators, Trade-Offs, Current Trends, Common Drawbacks, and Novel Solution Proposals," Aug. 01, 2022, *MDPI*. doi: 10.3390/math10152754.
- [29] *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON) : proceedings : Technology, Knowledge, and Society : 17-20*

- October 2019, Grand Hyatt Kochi Bolgatti, Kerala, India. IEEE, 2019.
- [30] 2019 International Conference on Computer and Information Sciences (ICCIS) : Jouf University - Aljouf - kingdom of Saudi Arabia, 03-04 April 2019. IEEE, 2019.
- [31] A. Imteaj, M. Hadi Amini, and P. M. Pardalos, "Toward Smart Contract and Consensus Mechanisms of Blockchain," 2021, pp. 15–28. doi: 10.1007/978-3-030-75025-1_2.
- [32] N. Szabo, "Formalizing and Securing Relationships on Public Networks," *First Monday*, vol. 2, no. 9, Sep. 1997, doi: 10.5210/fm.v2i9.548.
- [33] S. N. Khan, F. Loukil, C. Ghedira-Guegan, E. Benkhelifa, and A. Bani-Hani, "Blockchain smart contracts: Applications, challenges, and future trends," *Peer Peer Netw Appl*, vol. 14, no. 5, pp. 2901–2925, Sep. 2021, doi: 10.1007/s12083-021-01127-0.
- [34] K. L. Jones, "Blockchain: Building Consensus and Trust across the Space Sector," 2019.
- [35] C. Cachin and M. Vukolić, "Blockchain consensus protocols in the wild," in *Leibniz International Proceedings in Informatics, LIPIcs*, Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, Oct. 2017. doi: 10.4230/LIPIcs.DISC.2017.1.
- [36] Giang-Truong Nguyen and Kyungbaek Kim, "A Survey about Consensus Algorithms Used in Blockchain," vol. 14, pp. 101–128, Feb. 2018, Accessed: Dec. 12, 2024. [Online]. Available: <https://jips-k.org/full-text/view?doi=10.3745/JIPS.01.0024>
- [37] S. Zhang and J.-H. Lee, "Analysis of the main consensus protocols of blockchain," *ICT Express*, vol. 6, no. 2, pp. 93–97, Jun. 2020, doi: 10.1016/j.icte.2019.08.001.
- [38] L. S. Sankar, M. Sindhu, and M. Sethumadhavan, "Survey of consensus protocols on blockchain applications," in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, Jan. 2017, pp. 1–5. doi: 10.1109/ICACCS.2017.8014672.
- [39] S. Verma, D. Yadav, and G. Chandra, "Introduction of Formal Methods in Blockchain Consensus Mechanism and Its Associated Protocols," *IEEE Access*, vol. 10, pp. 66611–66624, 2022, doi: 10.1109/ACCESS.2022.3184799.
- [40] C. Cachin and M. Vukolić, "Blockchain consensus protocols in the wild," in *Leibniz International Proceedings in Informatics, LIPIcs*, Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, Oct. 2017. doi: 10.4230/LIPIcs.DISC.2017.1.
- [41] M. S. Ferdous *et al.*, "Blockchain Consensus Algorithms: A Survey", doi: 10.48550/arXiv.2001.07091.
- [42] B. Bellaj, A. Ouaddah, E. Bertin, N. Crespi, and A. Mezrioui, "SOK: A Comprehensive Survey on Distributed Ledger Technologies," in *2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, IEEE, May 2022, pp. 1–16. doi: 10.1109/ICBC54727.2022.9805533.
- [43] M. Salimitari and M. Chatterjee, "A Survey on Consensus Protocols in Blockchain for IoT Networks," Sep. 2018, [Online]. Available: <http://arxiv.org/abs/1809.05613>
- [44] "journal_jips_JIPS-2018-14-1-101 (1)".
- [45] O. Dib, A. Durand, K.-L. Brousmiche, E. Thea, and B. Hamida, "Consortium Blockchains: Overview, Applications and Challenges," 2018. [Online]. Available: <http://www.ariajournals.org/telecommunications/2018>,
- [46] Y. Zou, T. Meng, P. Zhang, W. Zhang, and H. Li, "Focus on Blockchain: A Comprehensive Survey on Academic and Application," *IEEE Access*, vol. 8, pp. 187182–187201, 2020, doi: 10.1109/ACCESS.2020.3030491.
- [47] F. Casino, T. K. Dasaklis, and C. Patsakis, "A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telematics and Informatics*, vol. 36, pp. 55–81, Mar. 2019, doi: 10.1016/j.tele.2018.11.006.
- [48] S. Thakur and V. Kulkarni, "Blockchain and Its Applications – A Detailed Survey," *Int J Comput Appl*, vol. 180, no. 3, pp. 29–35, Dec. 2017, doi: 10.5120/ijca2017915994.
- [49] H. Rathore, A. Mohamed, and M. Guizani, "A Survey of Blockchain Enabled Cyber-Physical Systems," *Sensors*, vol. 20, no. 1, p. 282, Jan. 2020, doi: 10.3390/s20010282.
- [50] I. Makhdoom, M. Abolhasan, and W. Ni, "Blockchain for IoT: The Challenges and a Way Forward," *Scitepress*, Sep. 2018, pp. 594–605. doi: 10.5220/0006905605940605.
- [51] M. A. Parssinen, M. Kotila, R. Cuevas Rumin, A. Phansalkar, and J. Manner, "Is Blockchain Ready to Revolutionize Online Advertising?," *IEEE Access*, vol. 6, pp. 54884–54899, 2018, doi: 10.1109/ACCESS.2018.2872694.
- [52] N. Hanis, A. Rasid, and O. K. Blockchain, "Blockchain Technology in e-Voting: Comparative Study." [Online]. Available: <https://www.researchgate.net/publication/342196799>
- [53] V. Clincy and H. Shahriar, "Blockchain development platform comparison," in *Proceedings - International Computer Software and Applications Conference*, IEEE Computer Society, Jul. 2019, pp. 922–923. doi: 10.1109/COMPSAC.2019.00142.
- [54] J. Burdges *et al.*, "Overview of Polkadot and its Design Considerations," May 2020, [Online]. Available: <http://arxiv.org/abs/2005.13456>
- [55] H. Abbas, M. Caprolu, and R. Di Pietro, "Understanding Polkadot Through Graph Analysis: Transaction Model, Network Properties, and Insights," 2024, pp. 259–275. doi: 10.1007/978-3-031-47751-5_15.
- [56] H. Abbas, M. Caprolu, and R. Di Pietro, "Analysis of Polkadot: Architecture, Internals, and Contradictions," in *2022 IEEE International Conference on Blockchain (Blockchain)*, IEEE, Aug. 2022, pp. 61–70. doi: 10.1109/Blockchain55522.2022.00019.
- [57] G. Wang, J. Wang, L. Lai, and F. Yu, "Accountability and Forensics in Blockchains: XDC Consensus Engine DPoS 2.0," Aug. 2021, [Online]. Available: <http://arxiv.org/abs/2108.01420>
- [58] XinFin (XDC) Hybrid Blockchain R&D Team, "XinFin Network-XDC Consensus Algorithm," May 2021.
- [59] C. Cachin and M. Vukolić, "Blockchain Consensus Protocols in the Wild," Jul. 2017, [Online]. Available: <http://arxiv.org/abs/1707.01873>
- [60] L. Baresi, G. Quattrocchi, D. A. Tamburri, and L. Terracciano, "A declarative modelling framework for the deployment and management of blockchain applications," in *Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems*, New York, NY, USA: ACM, Oct. 2022, pp. 311–321. doi: 10.1145/3550355.3552417.
- [61] T. Chitra, M. Quaintance, S. Haber, and W. Martino, "Agent-Based Simulations of Blockchain protocols illustrated via Kadana's Chainweb," in *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, Jun. 2019, pp. 386–395. doi: 10.1109/EuroSPW.2019.00049.

Açık Kaynaklı Biyomedikal Sistemlerin Hastane Finansal Yönetimi Açısından Değerlendirilmesi

Evaluation of Open Source Biomedical Systems in Terms of Hospital Financial Management

Türker Togay Aksoy
Biomedical Device Technology
Anadolu BİL Vocational School
Istanbul Aydın University
Istanbul, Türkiye
turkeraksoy@aydin.edu.tr

Serkan Çizmeciogulları
Biomedical Device Technology
Vocational School of Technical
Sciences
Kırşehir Ahi Evran University
Kırşehir, Türkiye
serkan.cizmeciogullari@ahievran.edu.tr

Şafak Yüksel
Computer Programming
Anadolu BİL Vocational School
Istanbul Aydın University
Istanbul, Türkiye
safakyuksel@aydin.edu.tr

Özetçe— Açık kaynaklı biyomedikal sistemlerin hastane finansal yönetimine entegrasyonu, operasyonel verimliliğin artırılması, maliyetlerin azaltılması ve finansal şeffaflığın artırılması için dönüştürücü bir strateji olarak ortaya çıkmıştır. Bu çalışma, açık kaynaklı sistemlerin hastane operasyonlarının çeşitli yönleri üzerindeki etkisini incelemektedir. Açık kaynaklı biyomedikal sistemlerin uygulanmasının, hastane departmanları arasında kesintisiz veri entegrasyonunu kolaylaştırdığı, kritik bilgilere gerçek zamanlı erişim sağladığı ve iş akışlarını kolaylaştırdığı gösterilmiştir. Bu sistemler aynı zamanda sağlık profesyonelleri, araştırmacılar ve idari personel arasındaki işbirliğini de teşvik ederek operasyonel verimliliğin artmasına ve araştırma girişimlerinin hızlanmasına yol açar. Ayrıca, açık kaynaklı sistemlerin benimsenmesi hastanelerde önemli maliyet tasarruflarıyla ilişkilendirilmiştir. Bu sistemler, cerrahi prosedürleri optimize ederek, veri yönetimini geliştirerek ve idari giderleri azaltarak sağlık kurumlarının daha uygun maliyetli çalışmasına olanak sağlar. Açık kaynak platformlarının işbirlikçi doğası aynı zamanda yenilikçiliği ve sürekli iyileştirmeyi teşvik ederek maliyetlerin azaltılmasına daha fazla katkıda bulunur. Mali şeffaflığa ilişkin olarak, açık kaynaklı biyomedikal sistemler veri erişilebilirliğini, hesap verebilirliği ve paydaş katılımını artırır. Bu sistemler, mali bilgilerin kolayca erişilebilir olmasını sağlayarak şeffaflık ve hesap verebilirlik kültürünü teşvik ederek paydaşların mali uygulamaları incelemesine ve hastane yöneticilerini hesap verebilir kılmasına olanak tanır. Açık kaynak sistemleriyle ilişkili maliyet tasarrufları, finansal raporlama yeteneklerinin geliştirilmesine ve şeffaflık girişimlerine yatırım yapılmasına da yönlendirilebilir. Açık kaynaklı biyomedikal sistemlerin hastane finansal yönetiminde stratejik olarak uygulanması, sağlık kurumları için dönüştürücü bir fırsat sunmaktadır. Bu sistemler operasyonel verimliliği artırarak, maliyetleri düşürerek ve mali şeffaflığı artırarak hastane operasyonlarının sürdürülebilirliğine ve etkinliğine önemli ölçüde katkıda bulunabilir ve sonuç olarak daha iyi hasta bakımı ve çıktılar sağlayabilir.

Anahtar Kelimeler — Açık kaynaklı biyomedikal sistemler; Hastane finansal yönetimi; Operasyonel verimlilik, Maliyet düşüşü, Finansal şeffaflık

Abstract— The integration of open-source biomedical systems into hospital financial management has emerged as a transformative strategy for enhancing operational efficiency, reducing costs, and improving financial transparency. This study examines the impact of open-source systems on various

aspects of hospital operations. The implementation of open-source biomedical systems has been shown to facilitate seamless data integration across hospital departments, enabling real-time access to critical information and streamlining workflows. These systems also foster collaboration among healthcare professionals, researchers, and administrative staff, leading to improved operational efficiency and accelerated research initiatives. Furthermore, the adoption of open-source systems has been linked to significant cost savings in hospitals. By optimizing surgical procedures, enhancing data management, and reducing administrative expenses, these systems empower healthcare institutions to operate more cost-effectively. The collaborative nature of open-source platforms also promotes innovation and continuous improvement, further contributing to cost reductions. Regarding financial transparency, open-source biomedical systems enhance data accessibility, accountability, and stakeholder engagement. By making financial information readily available, these systems foster a culture of transparency and accountability, enabling stakeholders to scrutinize financial practices and hold hospital administrators accountable. The cost savings associated with open-source systems can also be redirected towards improving financial reporting capabilities and investing in transparency initiatives. The strategic implementation of open-source biomedical systems in hospital financial management presents a transformative opportunity for healthcare institutions. By improving operational efficiency, reducing costs, and enhancing financial transparency, these systems can significantly contribute to the sustainability and effectiveness of hospital operations, ultimately leading to better patient care and outcomes.

Keywords — Open-source biomedical systems; Hospital financial management; Operational efficiency; Cost reduction; Financial transparency

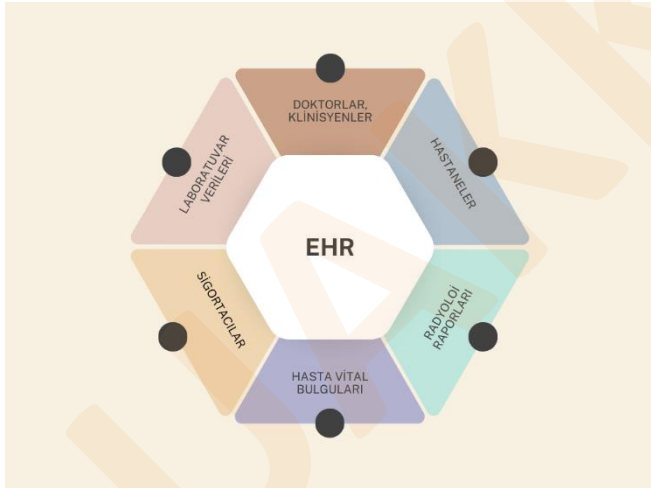
I. GİRİŞ

Açık kaynaklı biyomedikal sistemlerin hastane finansal yönetimine entegrasyonu, operasyonel verimliliğin artırılması, maliyetlerin azaltılması ve finansal şeffaflığın artırılması açısından giderek daha önemli bir strateji olarak kabul edilmektedir. Açık kaynaklı platformlar, hastanelere finansal yönetim araçlarını kendi özel ihtiyaçlarını karşılayacak şekilde özelleştirme esnekliği sunarak daha duyarlı ve uyarlanabilir bir finansal ortam teşvik eder. Bu

uyarlanabilirlik, hastanelerin hükümet tahsisleri, sigorta geri ödemeleri ve hasta ücretleri dahil olmak üzere karmaşık finansman kaynaklarına yönelmek zorunda kaldığı, gelişen sağlık hizmetleri finansman modelleri bağlamında özellikle önemlidir [1,2].

Açık kaynaklı sistemlerin başlıca avantajlarından biri, hastane içindeki çeşitli departmanlar arasında kapsamlı veri entegrasyonunu kolaylaştırma yetenekleridir. Bu sistemler, birden fazla kaynaktan gelen finansal verileri birleştirerek hastane yöneticilerinin finansal performansa ilişkin gerçek zamanlı bilgiler edinmesine, eğilimleri belirlemesine ve kaynak tahsisi ve bütçeleme konusunda bilinçli kararlar almasına olanak tanır [3]. Örneğin, Slovenya ve Hırvatistan'daki çalışmalarda vurgulandığı gibi maliyet muhasebesi sistemlerinin uygulanması, açık kaynaklı çözümlerinin finansal raporlamanın doğruluğunu ve etkinliğini nasıl artırabileceğini ve sonuçta daha iyi finansal yönetim uygulamalarına nasıl yol açabileceğini göstermektedir [3].

Ayrıca, açık kaynaklı sistemlerin kullanılması, özel mülk yazılım lisanslarıyla ilgili mali yükü önemli ölçüde azaltabilir. Hastaneler sıklıkla kısıtlı bütçelerle ve artan işletme maliyetleriyle karşı karşıyadır; dolayısıyla açık kaynaklı çözümlerinin benimsenmesi, hasta bakımı ve hastane altyapısının iyileştirilmesine yönlendirilebilecek önemli tasarruflara yol açabilir [4]. Açık kaynaklı elektronik sağlık kaydı (Electronic Health Record: EHR) projelerinin yönetimi, işbirlikçi gelişimin, maliyet etkinliğini korurken sağlık hizmeti sağlayıcılarının benzersiz ihtiyaçlarını karşılayan sağlam sistemlere yol açabileceğini göstermiştir [4]. EHR platformunu oluşturan öğeler Şekil 1'de gösterilmiştir.



Şekil 1. EHR platformunu oluşturan öğeler

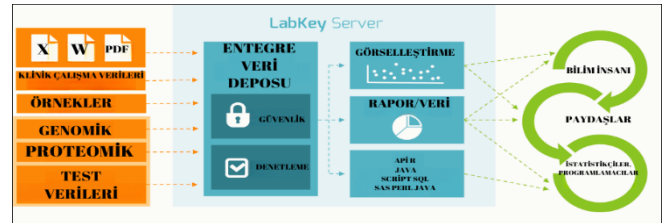
Açık kaynak sistemleri, maliyet tasarruflarının yanı sıra finansal yönetimde şeffaflığı ve hesap verebilirliği de destekler. Hastaneler, finansal verileri paydaşların erişimine sunarak güveni artırabilir ve kaynakların verimli bir şekilde kullanılmasını sağlayabilir. Bu şeffaflık, vergi mükelleflerine ve düzenleyici kurumlara karşı sorumluluğun çok önemli olduğu kamu sağlık hizmetleri ortamlarında özellikle önemlidir [5]. Ayrıca, raporlama araçlarını açık kaynak platformlarda özelleştirebilme yeteneği, hastanelerin düzenleyici gerekliliklere daha etkili bir şekilde uymasına olanak tanıyarak mali kötü yönetim riskini azaltır [6].

Açık kaynak sistemlerinin hastane mali yönetimi üzerindeki etkisi, özellikle COVID-19 salgınının yol açtığı son zorluklar bağlamında anlamlıdır. Pandemi, hızla değişen koşullara uyum sağlayabilecek çevik finansal yönetim uygulamalarına olan ihtiyacı artırdı. Açık kaynaklı çözümleri, bu zorluklara yanıt vermek için gerekli esnekliği sağlayarak hastanelerin finansal stratejilerini gerçek zamanlı olarak optimize etmelerine olanak sağlayabilir [7,8]. Örneğin, bir geri ödeme mekanizması olarak teşhisle ilgili grupların (DRG'ler) uygulamaya konulması, finansal yönetim uygulamalarında, klinik ve finansal verilerin entegrasyonuna izin veren açık kaynaklı sistemler tarafından etkili bir şekilde desteklenebilecek bir değişimi gerektirmiştir [9]. Sonuç olarak, hastane finansal yönetiminde açık kaynaklı biyomedikal sistemlerin benimsenmesi sağlık kurumları için dönüştürücü bir fırsat sunmaktadır. Veri entegrasyonunu geliştirerek, maliyetleri azaltarak, şeffaflığı teşvik ederek ve değişen koşullara uyum sağlamak için gereken esnekliği sağlayarak bu sistemler hastanelerin mali sağlığını önemli ölçüde iyileştirebilir. Sağlık hizmetleri ortamı gelişmeye devam ettikçe, hastane operasyonlarının sürdürülebilirliğini ve verimliliğini sağlamak için açık kaynaklı çözümlerin stratejik olarak uygulanması zorunlu olacaktır.

II. AÇIK KAYNAKLI BİYOMEDİKAL SİSTEMLERİN HASTANE YÖNETİMİNE KATKIDA BULUNAN BİLEŞENLERİ

A. Operasyonel Verimlilik

Açık kaynaklı biyomedikal sistemlerin hastanelerde uygulanmasının, sağlık hizmeti sunumunun çeşitli boyutlarında operasyonel verimliliği arttırdığı gösterilmiştir. Bu sistemler hastanelere bilgi teknolojisi çözümlerini belirli operasyonel ihtiyaçları karşılayacak şekilde özelleştirme ve uyarlama esnekliği sağlayarak iş akışını, veri yönetimini ve genel hasta bakımını iyileştirir. Bu çalışmada, veri entegrasyonu, işbirliği ve maliyet etkinliğine odaklanarak açık kaynaklı biyomedikal sistemlerin hastane operasyonları üzerindeki etkilerini araştırılmıştır. Açık kaynaklı sistemlerin dikkate değer avantajlarından biri, hastane içindeki farklı departmanlar arasında kesintisiz veri entegrasyonunu kolaylaştırma yetenekleridir. Örneğin Şekil 2'de görüldüğü üzere LabKey Server gibi sistemler, çeşitli biyomedikal verilerin entegrasyonuna olanak tanıyarak sağlık profesyonellerinin hasta bilgilerine, laboratuvar sonuçlarına ve klinik verilere gerçek zamanlı olarak erişmesine ve bunları analiz etmesine olanak tanır [10].



Şekil 2. LabKey Server platformu

Bu entegrasyon, kapsamlı veri kümelerine dayanarak bilinçli seçimler yapılabilmesini sağlayarak iş akışlarını kolaylaştırır ve karar verme süreçlerini geliştirir. Biolink Modeli, açık kaynaklı çerçevelerin, çeşitli biyomedikal uygulamalar

arasında birlikte çalışabilirliği teşvik ederek bilgi grafiklerinin uyumlaştırılmasını nasıl destekleyebileceğini örneklendirmektedir [11]. Bu tür bir birlikte çalışabilirlik, sağlık hizmeti sağlayıcılarının etkili bir şekilde işbirliği yapabilmesini ve kritik bilgileri hiçbir engel olmadan paylaşabilmesini sağlamak için çok önemlidir. Ayrıca, açık kaynaklı biyomedikal sistemler sağlık profesyonelleri, araştırmacılar ve idari personel arasındaki işbirliğini teşvik etmektedir. Hastaneler, işbirlikçi araştırma ve veri paylaşımını destekleyen platformları kullanarak operasyonel verimliliklerini artırabilir. Örneğin, Hizmet Seti Olarak Biyomedikal Araştırma Altyapı Yazılımı (BRISKit), açık kaynaklı çözümlerin katılımcı yönetiminden veri analizine kadar tüm araştırma ekosistemini nasıl yönetebileceğini ve böylece biyomedikal araştırmanın verimliliğini nasıl artırabileceğini göstermektedir [12]. Bu işbirlikçi ortam, araştırma girişimlerini hızlandırır ve bulguların hızla yaygınlaştırılması ve hasta bakımında uygulanabilmesi nedeniyle gelişmiş klinik uygulamalara dönüşür.

B. Maliyetlerdeki Düşüş

Hastanelerde açık kaynaklı biyomedikal sistemlerin benimsenmesi, bakım kalitesini artırırken operasyonel maliyetleri düşürmeye yönelik önemli bir strateji olarak ortaya çıkmıştır. Bu sistemler sağlık kurumlarına bilgi teknolojisi çözümlerini özelleştirme esnekliği sağlayarak çeşitli operasyonel alanlarda verimliliğin artmasına ve maliyet tasarrufuna yol açar. Bu çalışmada, açık kaynaklı biyomedikal sistemlerin hastane maliyetleri üzerindeki etkilerini araştırılmış ve bunların cerrahi prosedürler, veri yönetimi ve genel kaynak kullanımı üzerindeki etkilerine odaklanılmıştır. Açık kaynaklı sistemlerin maliyetin azaltılmasına katkıda bulunduğu en dikkat çekici alanlardan biri cerrahi prosedürlerdir. Çalışmalar, çoğunlukla açık kaynak teknolojilerle desteklenen minimal invaziv cerrahi tekniklerin, geleneksel açık ameliyatlara kıyasla önemli maliyet tasarrufları sağlayabileceğini göstermiştir. Örneğin, robot yardımlı laparoskopik prosedürlerin karşılaştırmalı bir analizi, bu tekniklerin yalnızca hastanede kalış süresini kısaltmakla kalmayıp aynı zamanda postoperatif komplikasyonları da en aza indirdiğini ve sonuçta genel hastane maliyetlerinin düşmesine yol açtığını göstermiştir [13,14]. Açık kaynaklı sistemler aracılığıyla cerrahi iş akışlarını kolaylaştırma yeteneği, ameliyathanelerin verimliliğini artırır; bu da maliyetlerin etkin bir şekilde yönetilmesi açısından çok önemlidir.

Ayrıca, açık kaynaklı elektronik sağlık kaydı (EHR) sistemlerinin entegrasyonu, daha iyi veri yönetimi ve kaynak tahsisini kolaylaştırır. Sağlık uzmanlarına hasta verilerine gerçek zamanlı erişim sağlayan bu sistemler, daha bilinçli karar alma olanağı sağlar, bu da kalış sürelerinin azalmasına ve hasta sonuçlarının iyileşmesine yol açabilir. Örneğin, açık kaynaklı EHR sistemleri tarafından desteklenen standartlaştırılmış perioperatif bakım yollarının uygulanması, daha kısa iyileşme süreleri ve daha az komplikasyon nedeniyle hastane maliyetlerinin azalmasıyla ilişkilendirilmiştir [15]. Hasta verilerini verimli bir şekilde takip etme ve analiz etme yeteneği, hastanelerin operasyonlarını optimize etmesine ve gereksiz harcamaları azaltılmasına olanak tanır. Cerrahi verimlilik ve veri yönetimine ek olarak açık kaynaklı sistemler, idari

maliyetleri azaltarak hastanelerin genel mali sağlığını önemli ölçüde etkileyebilir. Geleneksel tescilli sistemler genellikle yüksek lisans ücretleri ve bakım maliyetleriyle birlikte gelir ve bu da hastane bütçelerini zorlayabilir. Bunun aksine, açık kaynaklı çözümler bu masrafları ortadan kaldırarak hastanelerin kaynakları daha etkili bir şekilde tahsis etmesine olanak tanır. Bir çalışma, gelişmekte olan ülkelerde açık kaynaklı hastane bilgi sistemlerinin (HIS) kullanımının fizibilite ve maliyet etkinliği gösterdiğini ve sağlık hizmeti sunumunda önemli tasarruf potansiyelini ortaya koyduğunu vurgulamıştır [16]. Hastaneler, açık kaynak sistemlerini benimseyerek personel eğitimi ve altyapı iyileştirmeleri gibi diğer kritik alanlara yatırım yaparak operasyonel verimliliği daha da artırabilir.

Ayrıca, açık kaynaklı sistemlerin işbirlikçi doğası, hastane operasyonlarında yenilikçiliği ve sürekli iyileştirmeyi teşvik eder. Hastaneler, en iyi uygulamaları paylaşabilir ve kendilerine özgü zorluklara yanıt veren özelleştirilmiş çözümler geliştirerek verimliliğin artmasına ve maliyet tasarrufuna yol açabilir. Örneğin, kronik hastalıkları yönetmek için açık kaynaklı platformların kullanılması, hastaneye yatışlarla ilişkili maliyetleri kontrol ederken hasta yönetimi üzerinde de olumlu etkiler göstermiştir [17]. Bu işbirlikçi yaklaşım yalnızca maliyetleri düşürmekle kalmayıp, aynı zamanda hastalara sağlanan bakımın kalitesini de artırmaktadır. Sonuç olarak, açık kaynaklı biyomedikal sistemlerin hastanelere entegrasyonu, bakım kalitesini artırırken operasyonel maliyetleri azaltmak için dönüştürücü bir fırsat sunmaktadır. Cerrahi verimliliği artırarak, veri yönetimini optimize ederek ve idari giderleri en aza indirerek bu sistemler, sağlık kurumlarının giderek daha karmaşık hale gelen sağlık hizmetleri ortamında daha etkili bir şekilde çalışmalarını sağlar. Hastaneler mali baskılarla karşı karşıya kalmaya devam ettikçe, açık kaynak teknolojilerinin stratejik olarak benimsenmesi, sürdürülebilir maliyet düşüşleri sağlamak ve hasta sonuçlarını iyileştirmek için gerekli olacaktır.

C. Finansal Şeffaflık

Açık kaynaklı biyomedikal sistemlerin hastanelere entegrasyonunun, etkili sağlık yönetiminin kritik bir yönü olan finansal şeffaflık açısından önemli etkileri vardır. Mali şeffaflık, bir kuruluşun açık raporlama, erişilebilir veriler ve karar alma süreçlerinde açıklık dahil olmak üzere mali bilgilerini paylaştığı açıklık derecesini ifade eder [18]. Bu çalışmada, veri erişilebilirliğine, hesap verebilirliğe ve paydaş katılımına odaklanarak açık kaynak sistemlerinin hastanelerde finansal şeffaflığı nasıl geliştirdiği araştırılmıştır. Açık kaynaklı biyomedikal sistemlerin finansal şeffaflığı artırmasının başlıca yollarından biri, gelişmiş veri erişilebilirliğidir. Bu sistemler, hastanelerin finansal verileri kullanıcı dostu bir şekilde yönetmesine ve paylaşmasına olanak tanıyarak hastalar, düzenleyici kurumlar ve kamu dahil olmak üzere paydaşların ilgili finansal bilgilere erişmesini kolaylaştırır. Örneğin, açık kaynaklı elektronik sağlık kaydı (EHR) sistemleri, klinik bilgilerin yanı sıra finansal verileri de içerecek şekilde tasarlanabilir ve hastane operasyonlarının kapsamlı bir görünümünü sağlayabilir. Bu entegrasyon, paydaşların halihazırda mevcut verilere dayanarak kurumun mali durumunu değerlendirebilmeleri nedeniyle bilinçli karar vermelerini kolaylaştırır.

Ayrıca, açık kaynaklı sistemler hastanenin mali yönetiminde hesap verebilirliği teşvik eder. Hastaneler, finansal verileri kamuya açık hale getirerek, finansal kararların hem iç hem de dış paydaşlar tarafından incelendiği bir hesap verebilirlik kültürünü teşvik edebilir. Bu şeffaflık, paydaşların hastane yöneticilerini mali uygulamalarından sorumlu tutabilmesi nedeniyle mali kötü yönetim ve yolsuzlukla ilgili sorunların azaltılmasına yardımcı olabilir. Örneğin, hastanelerde fiyat şeffaflığı düzenlemelerinin uygulanmasının, uyumu teşvik ettiği ve finansal raporlama uygulamalarını geliştirerek sağlık harcamalarında daha fazla hesap verebilirliğe yol açtığı gösterilmiştir [19,20]. Açık kaynaklı sistemler, doğru ve zamanında raporlama için gerekli altyapıyı sağlayarak bu tür düzenlemelere uyumu kolaylaştırabilir. Açık kaynaklı biyomedikal sistemlerin bir diğer önemli faydası da paydaş katılımını artırma yetenekleridir. Hastaneler, şeffaf finansal bilgiler sağlayarak hastalarla ve toplumla daha etkili bir şekilde etkileşime geçebilir. Bu katılım, güven oluşturma ve paydaşların sağlık hizmetlerinin finansmanı ve kaynak tahsisi hakkındaki tartışmalara katılma konusunda kendilerini yetkili hissettikleri işbirlikçi bir ortamın teşvik edilmesi açısından çok önemlidir.

Örneğin, sağlık hizmeti sağlayıcıları ile sektör paydaşları arasındaki mali bağları açığa çıkaran açık veri girişimleri, potansiyel çıkar çatışmalarının belirlenmesine yardımcı olarak daha şeffaf ve hesap verebilir bir sağlık sistemi teşvik edebilir [21]. Bu tür girişimler aynı zamanda mali çatışmaları ele almak ve sağlık sektöründe genel şeffaflığı artırmak için hastaneler, profesyonel topluluklar ve devlet kurumları arasındaki iş birliğini de teşvik edebilir. Ayrıca, açık kaynaklı sistemlerin kullanılması maliyet tasarruflarına yol açabilir ve bu da finansal şeffaflık girişimlerinin iyileştirilmesine yönlendirilebilir. Hastaneler, özel yazılımla ilgili lisans ücretlerini ortadan kaldırarak, mali raporlama yeteneklerini geliştirmek için kaynak tahsis edebilir ve mali verileri etkili bir şekilde yönetmek için eğitim personeline yatırım yapabilir. Finansal şeffaflığa yapılan bu yatırım sonuçta daha iyi finansal yönetim uygulamalarına, daha iyi hasta bakımına ve gelişmiş operasyonel verimliliğe yol açabilir.

Sonuç olarak, hastanelerde açık kaynaklı biyomedikal sistemlerin benimsenmesi finansal şeffaflığın artırılmasında hayati bir rol oynamaktadır. Bu sistemler, veri erişilebilirliğini iyileştirerek, hesap verebilirliği teşvik ederek ve paydaş katılımını teşvik ederek sağlık kurumlarının daha şeffaf ve verimli çalışmasını sağlar. Sağlık hizmetleri ortamı gelişmeye devam ettikçe, sürdürülebilir mali şeffaflığın sağlanması ve hastalara sağlanan bakımın genel kalitesinin iyileştirilmesi için açık kaynak teknolojilerinin stratejik olarak uygulanması zorunlu olacaktır.

III. TARTIŞMA

Hastanelerde operasyonel verimlilik, maliyetlerin azaltılması ve finansal şeffaflığa ilişkin tartışma, artan sağlık hizmetleri maliyetleri ve sürdürülebilir sağlık hizmeti sunum sistemlerine olan ihtiyaç bağlamında giderek daha fazla önem kazanmaktadır. Hastaneler, genellikle idari ve genel maliyetlerin hakim olduğu ve toplam işletme giderlerinin yaklaşık %21,4'ünü oluşturan maliyet yapılarını yönetmede önemli zorluklarla karşı karşıyadır [22]. Bu, hastanelerin operasyonel verimliliği artırırken aynı zamanda gereksiz harcamaları azaltabilecek stratejik maliyet yönetimi uygulamalarını benimsemelerinin gerekliliğini

vurgulamaktadır. Operasyonel verimliliği artırmaya yönelik etkili bir yaklaşım, faaliyet bazlı maliyetleme (Activity-Based Costing: ABC) ve hedef maliyetleme gibi ileri yönetim muhasebesi tekniklerinin uygulanmasıdır. Bu metodolojiler, hastanelerin maliyet etkenlerini daha net anlamalarına ve verimliliğin gerçekleştirilebileceği alanları belirlemelerine olanak tanır [23]. Örneğin, hastaneler belirli hasta bakım hizmetleriyle ilgili maliyetleri analiz ederek kaynak tahsisi ve hizmet sunumu konusunda bilinçli kararlar alabilir ve sonuçta operasyonel maliyetlerin azalmasına yol açabilir. ABC'nin uygulanması aynı zamanda daha iyi fiyatlandırma stratejilerini kolaylaştırabilir ve hastanelerin kaliteli bakımı korurken rekabetçi kalmasını sağlayabilir.

Ayrıca hasta hacmindeki dalgalanmalara rağmen belirli maliyetlerin sabit kalma eğilimini ifade eden "yapışkan maliyetler" (Sticky Costs) kavramı, hastane maliyet davranışını anlamak için çok önemlidir [24]. Hastaneler bütçeleri planlarken ve finansal performansı tahmin ederken bu dinamiklerin farkında olmalıdır. Hastane yöneticileri, hangi maliyetlerin değişken, hangilerinin sabit olduğunu tanıyarak hasta talebi ve operasyonel kapasiteyle uyumlu daha etkili maliyet kontrol önlemleri uygulayabilir. Finansal şeffaflık, hastane yönetiminin operasyonel verimliliği artırabilecek ve maliyetleri azaltabilecek bir diğer kritik yönüdür. Sağlık hizmetlerinde son zamanlarda fiyat şeffaflığı yönündeki baskının hastaneler üzerinde önemli etkileri vardır. Hastaneler, fiyatlandırma bilgilerinin hastalar için erişilebilir olmasını sağlamak, uygun maliyetli bakım seçeneklerini teşvik eden rekabetçi bir ortamı teşvik edebilir [25]. Ancak gerçek mali şeffaflığa ulaşmak, düzenlemelere uyumdan daha fazlasını gerektirir; finansal konularda açık iletişime ve hesap verebilirliğe öncelik vermek için kuruluşlar içinde kültürel bir değişimi gerektirmektedir. Şeffaflığın yanı sıra hastaneler, operasyonel verimliliği artırmak ve maliyetleri azaltmak için teknolojiden yararlanabilir. Elektronik sağlık kayıtlarının (EHR'ler) ve veri analizi araçlarının entegrasyonu idari süreçleri kolaylaştırabilir, evrak işlerini azaltabilir ve hasta bakımı koordinasyonunu geliştirebilir [22]. Bu teknolojiler yalnızca operasyonel iş akışlarını iyileştirmekle kalmıyor, aynı zamanda maliyet yönetimi ve kaynak kullanımına ilişkin değerli bilgiler sağlayarak hastanelerin finansal performansı artıran veriye dayalı kararlar almasına olanak tanıyor.

Ayrıca, COVID-19 salgını hastane operasyonlarında uyum sağlamanın öneminin altını çizdi. Kriz, birçok hastanenin maliyet yapılarını ve operasyonel stratejilerini yeniden değerlendirmesine yol açarak bakıma erişimi korurken maliyetleri azaltabilen tele-sağlık hizmetleri gibi bakım sunumunda yeniliklere yol açtı [22]. Hastaneler pandemi sonrası ortamda ilerlemeye devam ettikçe operasyonel verimlilik ve maliyet yönetimine ilişkin öğrenilen dersler uzun vadeli sürdürülebilirlik açısından hayati önem taşıyacaktır. Sonuç olarak hastanelerde operasyonel verimliliğin artırılması, maliyetlerin azaltılması ve finansal şeffaflığın artırılması, birbirleriyle bağlantılı ve çok yönlü bir yaklaşım gerektiren hedeflerdir. Hastaneler, gelişmiş maliyet yönetimi tekniklerini benimseyerek, finansal şeffaflığı benimseyerek, teknolojiden yararlanarak ve değişen koşullara uyum sağlayabilerek, giderek daha karmaşık hale gelen sağlık hizmetleri ortamında kendilerini başarı için konumlandırabilirler. Bu ilkelere sürekli bağlılık yalnızca

hastane operasyonlarına fayda sağlamakla kalmayacak, aynı zamanda hasta bakımını ve sonuçlarını da iyileştirecektir.

IV. SONUÇ

Çağdaş sağlık hizmetleri ortamında, hastanelerde operasyonel verimliliğin artırılması, maliyetlerin azaltılması ve mali şeffaflığın artırılması, hasta bakımını ve kurumsal sürdürülebilirliği önemli ölçüde etkileyebilecek kritik hedeflerdir. Bu hedeflere ulaşmak için stratejik yönetim uygulamalarını, yenilikçi bakım modellerini ve teknolojinin etkin kullanımını içeren çok yönlü bir yaklaşım esastır.

Operasyonel verimliliği artırmaya yönelik temel stratejilerden biri, birinci basamak hekimlerini ve uzmanlarını içeren işbirlikçi bakım modellerinin uygulanmasıdır. Kanıtlar, bu tür işbirliklerinin gereksiz hastane ziyaretlerinin azalmasına ve sonuçta operasyonel maliyetlerin azalmasına yol açabileceğini göstermektedir [26]. Ayrıca, multidisipliner bakım ekiplerinin entegrasyonu, hastanede kalış sürelerini azaltabilen ve hasta sonuçlarını iyileştirebilen, dolayısıyla maliyetlerin azaltılmasına katkıda bulunabilen gelişmiş hastalık yönetimi ile ilişkilendirilmiştir [27]. Bu yaklaşım, genellikle yüksek sağlık hizmetleri maliyetleriyle ilişkilendirilen kronik durumların yönetiminde koordineli bakımın önemini vurgulayan bulgularla uyumludur. Ek olarak, yatarak tedavi hizmetlerinin yerine ayakta tedavi hizmetlerinin ikame edilmesi, genel sağlık harcamalarının azaltılması konusunda ümit verici olmuştur. Sistematik incelemeler, özellikle hasta hacimleri yüksek olduğunda ayakta tedavi bakımının geleneksel hastane ortamlarına göre daha uygun maliyetli olabileceğini göstermektedir [28]. Bu değişim, yalnızca hastanelerin üzerindeki mali yükü hafifletmekle kalmıyor, aynı zamanda hastaların bakıma erişimini de geliştirerek genel sağlık sonuçlarını iyileştirmektedir.

Finansal şeffaflık, hastanelerin güven ve hesap verebilirliği geliştirmek için ele alması gereken bir diğer önemli husustur. Hastaların sağlık hizmetleri seçenekleriyle ilgili bilinçli kararlar almasına yardımcı olmak için fiyat şeffaflığını artırmayı amaçlayan son düzenlemeler uygulanmıştır [19,20]. Ancak hastaneler fiyatlandırma yapılarının ve geri ödeme oranlarının karmaşıklığı içindeyken bu bilgilerin hastalar için erişilebilir ve anlamlı olmasını sağlama konusunda zorluklar devam etmektedir [29]. Fiyatlandırma bilgilerinin etkili bir şekilde iletilmesi, hastaları güçlendirebilir ve potansiyel olarak sağlık hizmeti sağlayıcıları arasındaki rekabeti artırarak hastanelerin daha iyi finansal performansa ulaşmasını sağlayabilir. Ayrıca, hastane operasyonlarının maliyet liderliği veya farklılaşma stratejileriyle stratejik olarak hizalanması, finansal sonuçları önemli ölçüde etkileyebilir. Araştırmalar, tek bir stratejik odağı başarılı bir şekilde takip eden hastanelerin daha iyi finansal performans elde etme eğiliminde olduğunu göstermektedir [30]. Bununla birlikte, her iki stratejiyi de dengeleyebilenler kendilerini daha avantajlı bir konumda bulabilirler; bu da maliyet verimliliğini korurken pazar taleplerine yanıt vermede daha fazla esnekliğe olanak tanır. Teknolojinin operasyonel verimliliği artırmadaki rolü abartılamaz. Gelişmiş bilgi sistemlerinin benimsenmesi, kaynak tahsisini optimize edebilen ve israfı azaltabilen gerçek zamanlı veri yakalama ve analizini kolaylaştırır [31]. Ek olarak, hızlı iyileşme protokollerini destekleyen teknolojilerin kalış sürelerini ve ilgili maliyetleri azalttığı,

hastane ortamlarında verimliliği artıracak yenilikçi uygulamaların potansiyelini ortaya koyduğu gösterilmiştir [32].

Sonuç olarak hastanelerde operasyonel verimliliğin artırılması, maliyetlerin azaltılması ve finansal şeffaflığın sağlanması kapsamlı ve bütünlük bir yaklaşım gerektirmektedir. Hastaneler, işbirlikçi bakım modellerini teşvik ederek, ayakta tedavi hizmetlerini benimseyerek, fiyat şeffaflığını artırarak, operasyonları stratejik olarak hizalayarak ve teknolojiye yararlanarak yalnızca mali durumlarını iyileştirmekle kalmaz, aynı zamanda hastalara sağlanan bakımın kalitesini de artırabilir. Bu çabalar, modern sağlık hizmeti ortamının karmaşıklıklarının üstesinden gelmek ve sağlık kurumlarının sürdürülebilirliğini sağlamak için gereklidir.

KAYNAKLAR

- [1] E. Barasa, A. Manyara, S. Molyneux, and B. Tsofa, "Recentralization within decentralization: county hospital autonomy under devolution in Kenya," *PLOS ONE*, vol. 12, no. 8, p. e0182440, 2017. doi: 10.1371/journal.pone.0182440.
- [2] E. Barasa, S. Molyneux, M. English, and S. Cleary, "Setting healthcare priorities in hospitals: a review of empirical studies," *Health Policy and Planning*, vol. 30, no. 3, pp. 386–396, 2014. doi: 10.1093/heapol/czu010.
- [3] T. Jovanović, I. Dražić-Lutitsky, and D. Vašiček, "Implementation of cost accounting as the economic pillar of management accounting systems in public hospitals – the case of Slovenia and Croatia," *Economic Research-Ekonomska Istraživanja*, vol. 32, no. 1, pp. 3754–3772, 2019. doi: 10.1080/1331677x.2019.1675079.
- [4] P. Poba-Nzaou, J. Marsan, G. Paré, and L. Raymond, "Governance of open source electronic health record projects: a successful case of a hybrid model," in *Proceedings of HICSS*, pp. 2798–2807, 2014. doi: 10.1109/hicss.2014.350.
- [5] A. Millen and A. Stacey, "Financial literacy in South African healthcare professionals: an unmet need in health professions education," *South African Journal of Higher Education*, vol. 36, no. 3, pp. 123–142, 2022. doi: 10.20853/36-3-4647.
- [6] K. Kumar, A. Singh, K. James, L. McDougal, and A. Raj, "Gender bias in hospitalization financing from borrowings, selling of assets, contribution from relatives or friends in India," *Social Science & Medicine*, vol. 260, p. 113222, 2020. doi: 10.1016/j.socscimed.2020.113222.
- [7] J. Rhodes, T. Santos, and G. Young, "The early impact of the COVID-19 pandemic on hospital finances," *Journal of Healthcare Management*, vol. 68, no. 1, pp. 38–55, 2023. doi: 10.1097/jhm-d-22-00037.
- [8] A. Ahangar, "COVID-19 and reimbursement mechanisms in medical economics; retrospective (FFS) or prospective (DRG) payment healthcare systems, which one?," *American Journal of Biomedical Science & Research*, vol. 14, no. 2, pp. 167–173, 2021. doi: 10.34297/ajbsr.2021.14.001973.
- [9] W. Eiff, S. Schüring, B. Greitemann, and M. Karoff, "Redia – Auswirkungen der DRG-Einführung auf die Rehabilitation," *Die Rehabilitation*, vol. 50, no. 4, pp. 214–221, 2011. doi: 10.1055/s-0031-1275720.
- [10] E. Nelson, B. Piehler, J. Eckels, A. Rauch, M. Bellew, P. Hussey, and M. Igra, "Labkey server: an open source platform for scientific data integration, analysis and collaboration," *BMC Bioinformatics*, vol. 12, no. 1, 2011. doi: 10.1186/1471-2105-12-71.
- [11] D. Unni, S. Moxon, M. Bada, M. Brush, R. Bruskiwich, J. Caufield, and T. Zisk, "Bioblink model: a universal schema for knowledge graphs in clinical, biomedical, and translational science," *Clinical and Translational Science*, vol. 15, no. 8, pp. 1848–1855, 2022. doi: 10.1111/cts.13302.
- [12] O. Butters, S. Issa, J. Lusted, M. Newbury, R. Parsloe, N. Holden, and J. Tedds, "The biomedical research infrastructure software as a service kit (BRISKit): technical description," *F1000Research*, vol. 5, p. 1905, 2016. doi: 10.12688/f1000research.8736.1.
- [13] N. Aaronson, P. Neubauer, and B. Judson, "Cost effectiveness of transoral robotic surgery for the treatment of oropharyngeal squamous

- cell carcinoma: a systematic review," *Robotic Surgery Research and Reviews*, vol. 59, 2015. doi: 10.2147/rsrr.s64869.
- [14] E. Baker, S. Ross, R. Seshadri, R. Swan, D. Iannitti, D. Vrochides, and J. Martinie, "Robotic pancreaticoduodenectomy: comparison of complications and cost to the open approach," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 12, no. 3, pp. 554–560, 2015. doi: 10.1002/rcs.1688.
- [15] A. Page, F. Gani, K. Crowley, K. Lee, M. Grant, T. Zavadsky, and T. Pawlik, "Patient outcomes and provider perceptions following implementation of a standardized perioperative care pathway for open liver resection," *British Journal of Surgery*, vol. 103, no. 5, pp. 564–571, 2016. doi: 10.1002/bjs.10087.
- [16] C. Bagayoko, J. Dufour, S. Chaacho, O. Bouhaddou, and M. Fieschi, "Open source challenges for hospital information system (HIS) in developing countries: a pilot project in Mali," *BMC Medical Informatics and Decision Making*, vol. 10, no. 1, 2010. doi: 10.1186/1472-6947-10-22.
- [17] S. Purkayastha, R. Allam, P. Maity, and J. Gichoya, "Comparison of open-source electronic health record systems based on functional and user performance criteria," *Healthcare Informatics Research*, vol. 25, no. 2, p. 89, 2019. doi: 10.4258/hir.2019.25.2.89
- [18] S. RA, "The effect of transparency, performance evaluation, and financial reporting on achieving budget targets," *International Journal of Advanced Multidisciplinary*, vol. 2, no. 3, pp. 828–839, 2023. doi: 10.38035/ijam.v2i3.396.
- [19] T. D'Amore, G. Goh, P. Courtney, and G. Klein, "Do new hospital price transparency regulations reflect value in arthroplasty?," *Journal of the American Academy of Orthopaedic Surgeons*, 2022. doi: 10.5435/jaaos-d-21-00696.
- [20] E. Kong and Y. Ji, "Provision of hospital price information after increases in financial penalties for failure to comply with a US federal hospital price transparency rule," *JAMA Network Open*, vol. 6, no. 6, p. e2320694, 2023. doi: 10.1001/jamanetworkopen.2023.20694.
- [21] J. Mejía, A. Mejía, and F. Pestilli, "Open data on industry payments to healthcare providers reveal potential hidden costs to the public," *Nature Communications*, vol. 10, no. 1, 2019. doi: 10.1038/s41467-019-12317-z.
- [22] G. Bai and H. Zare, "Hospital cost structure and the implications on cost management during COVID-19," *Journal of General Internal Medicine*, vol. 35, no. 9, pp. 2807–2809, 2020. doi: 10.1007/s11606-020-05996-8.
- [23] R. Cooper and R. Slagmulder, *Target Costing and Value Engineering*, 2017. doi: 10.1201/9780203737378.
- [24] R. Balakrishnan, E. Labro, and N. Soderstrom, "Cost structure and sticky costs," *Journal of Management Accounting Research*, vol. 26, no. 2, pp. 91–116, 2014. doi: 10.2308/jmar-50831.
- [25] J. Spencer, "The three most important words in faculty workload: transparency, transparency, transparency," *The Department Chair*, vol. 34, no. 1, pp. 1–3, 2023. doi: 10.1002/dch.30522.
- [26] E. Wackers, S. Dulmen, B. Berden, J. Kremer, N. Stadhouders, and P. Jeurissen, "Improving performance in complex surroundings: a mixed methods evaluation of two hospital strategies in the Netherlands," *International Journal of Health Policy and Management*, vol. 12, p. 7243, 2023. doi: 10.34172/ijhpm.2023.7243.
- [27] J. Darba and A. Marsa, "Patient characteristics and hospitalisation costs of spinal muscular atrophy in Spain: a retrospective multicentre database analysis," *BMJ Open*, vol. 9, no. 11, p. e031271, 2019.
- [28] S. Hoof, T. Quanjel, M. Kroese, M. Spreeuwenberg, and D. Ruwaard, "Substitution of outpatient hospital care with specialist care in the primary care setting: a systematic review on quality of care, health and costs," *PLOS ONE*, vol. 14, no. 8, p. e0219957, 2019. doi: 10.1371/journal.pone.0219957.
- [29] Z. G. Gul, D. R. Sharbaugh, C. J. Guercio, et al., "Large variations in the prices of urologic procedures at academic medical centers 1 year after implementation of the Price Transparency Final Rule," *JAMA Network Open*, vol. 6, no. 1, p. e2249581, 2023. doi: 10.1001/jamanetworkopen.2022.49581.
- [30] A. Ghiasi, R. Weech-Maldonado, F. Zengul, and N. Puro, "The association between strategic group membership and hospital financial performance," *Journal of Hospital Management and Health Policy*, vol. 6, pp. 24–24, 2022. doi: 10.21037/jhmhp-21-71.
- [31] E. Landaas, G. Baird, R. Hansen, D. Flum, and S. Sullivan, "Integrating formal technology assessment into an integrated healthcare delivery system: smart innovation," *International Journal of Technology Assessment in Health Care*, vol. 36, no. 1, pp. 58–63, 2020. doi: 10.1017/s0266462319003465.
- [32] M. Lima, G. Borgomoni, and O. Mejía, "Teams, rapid recovery protocols and technology to resume cardiac surgery in the COVID-19 era," *Brazilian Journal of Cardiovascular Surgery*, vol. 36, no. 6, 2021. doi: 10.21470/1678-9741-2021-0035.

Review of End-to-End Data Analytics Platforms Streamlining Data Management and Decision-Making in Higher Education

Menekse Adar
Engineering and Computer Science
Engineering Management and
Technology
Chattanooga, United State
menekseberkeradar@gmail.com

Zachary Ridder
Engineering and Computer Science
Engineering Management and
Technology
Chattanooga, United State
zachary-ridder@utc.edu

Serkan Varol
Engineering and Computer Science
Engineering Management and
Technology
Chattanooga, United State
serkan-varol@utc.edu

Abstract—Data analytics platforms such as Alteryx, Tableau Prep, RapidMiner, and other similar data preparation and analytics tools have revolutionized how businesses and individuals handle large volumes of data without needing advanced programming skills. Likewise, these platforms have become important tools within higher education settings. They are increasingly utilized to perform complex data collection, wrangling, transformation, analysis, and visualizations in ways that were once complex or time-consuming. This review paper highlights the importance of these platforms in a data-driven education world, emphasizing their impact on operational efficiency and informed decision-making across diverse academic fields. It also explores the differences between these platforms, offering a comparative analysis of their unique features and functionalities. Overall, these package programs optimize data management, enhance productivity, and broaden access to data analytics across diverse user groups, all of which are essential in today's data-driven world.

Keywords—Alteryx, Tableau Prep, RapidMiner, KNIME, SAS, review

I. INTRODUCTION

As the field of data analytics continues to expand, a wide range of products and services are available to post-secondary students for free while they are enrolled in a university. Platforms like Alteryx and SAS typically require a licensing fee but offer free access to students while they are in college. On the other hand, programs such as RapidMiner and KNIME are available for anyone to download and use at no cost. However, while there are commonalities between many programs, such as the ability to import data from standard file types such as comma-separated values (CSV) or Microsoft Excel (XLSX), there are also many differences, such as minimum system requirements or if they require code to be written in order to process data. This paper will examine five data analytics platforms (Alteryx, Tableau, KNIME, RapidMiner, and SAS) that university students have access to for free during the period of their enrollment and compare their functionality across multiple dimensions.

II. COMPARISON OF SOFTWARE SETUP PROCESSES

The setup process is critical for users, and inadequate resources can make it complicated and time-consuming. Well-prepared and regularly updated documentation and tutorial videos explaining each step in detail are essential for a seamless setup experience. This section analyzes the setup processes of Alteryx, Tableau Prep, RapidMiner, KNIME, and SAS.

A. Alteryx

Alteryx is available only for Windows users, which presents a limitation for those using Linux or macOS. However, this can be mitigated by employing a Windows virtual machine or compatibility layers to enable the use of Alteryx on these systems.

Alteryx provides well-designed documentation and a dedicated YouTube channel featuring comprehensive tutorials, simplifying the setup process even for users with limited experience in software installation. These resources simplify the setup process, even for users unfamiliar with software installation. Despite its excellent documentation, Alteryx has higher system requirements than other tools. It requires a minimum of 500 GB of storage space and 16 GB of RAM, significantly higher than other platforms' requirements. [1]

B. Tableau Prep

Tableau offers a suite of applications, including Tableau Desktop, Server, and Prep. In this article, we focus on Tableau Prep. The company provides well-organized and easy-to-follow documentation for all its applications, which is especially helpful when managing multiple tools. However, users may occasionally find the setup process confusing, even with detailed instructions. Tableau also maintains a YouTube channel with exceptional content that helps users understand and utilize new features effectively. Tableau Prep has moderate hardware requirements, needing just 4 GB of RAM, which is considerably lower than Alteryx. Additionally, Tableau Prep supports macOS and virtual environments, making it more accessible to a broader range of users. [2]

C. RapidMiner

RapidMiner is a versatile tool that supports multiple operating systems, including Windows, Linux, and macOS, making it an excellent option for Linux users. The software requires at least 4 GB of RAM and more than 1 GB of free disk space, making it a lightweight option compared to Alteryx. RapidMiner provides well-prepared documentation tailored for each operating system, ensuring users can easily follow the setup process. Furthermore, the platform maintains a YouTube channel with tutorial videos, simplifying the installation and initial use for beginners. These resources make RapidMiner accessible and user-friendly for a wide range of users. [3]

D. KNIME

KNIME supports Windows, Ubuntu, and macOS, and it is free for personal use. Unlike some other tools, KNIME does not require a license key for setup, simplifying the process.

This advantage positions KNIME as a more convenient option than other tools. KNIME also provides well-designed documentation and a YouTube channel, ensuring users have access to resources that guide them effectively through the setup and usage process. [4]

space, which is on the higher end compared to other tools. The SAS interface appears somewhat outdated compared to its competitors. While SAS does not have its own YouTube channel, it provides users with comprehensive documentation. The use of a programming language instead of drag-and-drop modules may extend the learning curve for new users. [5]

E. SAS

SAS supports Windows and Linux but does not support macOS, which is similar to Alteryx. It requires 128 GB of disk

TABLE 1 Program Requirements and Licensing

Programs	Program Requirements and Licensing			
	OS Compatibility	Minimum Requirements	Opensource	License Type
Alteryx	Microsoft Windows Server 2012 64-bit or later Microsoft.NET Framework 4.6 or higher	Minimum Requirements Per Server Node CPU: Intel or AMD x86-64 (64-bit) Number of CPU Cores: 4 cores (8 logical cores) per node System Memory: 16 GB Recommended Minimum Disk Size: 500 GB	No	Students (1-year)
Tableau PREP	Windows MacOS Virtual Environments	Windows 8/8.1,10 (x64), 11 Intel Core i3 or AMD Ryzen 3 (Dual Core) 4GB memory or larger 2GB HDD free or larger CPUs must support SSE4.2 and POPCNT instruction sets macOS OS Big Sur 11.4+, macOS Monterey 12.6+ (for Tableau 2022.3+), macOS Ventura (for Tableau 2022.3+), macOS Sonoma (for Tableau 2022.3+); Apple Silicon machines require the use of macOS Ventura (13+) or newer Intel processors - Core i3 (Dual Core) or newer Apple Silicon processors (using Rosetta - 24.1 and below) Apple Silicon processors (version 24.2 or newer on MacOS Ventura or newer) 4GB memory or larger 2GB HDD free or larger	No	Students (1-year)
RapidMiner	Windows Linux MacOS	Windows 7, 8, 8.1, 10 (64-bit highly recommended) Linux (64-bit only) MacOS X 10.10 - 10.14 Dual core 2GHz processor 4GB RAM 1GB free disk space Screen resolution: 1280x1024	No	Yes (No limit)
KINME	Windows Windows Server Ubuntu MacOS	Windows 10, 11 Windows Server - 2016, 2019, 2022 Ubuntu 20.04 LTS and 22.04 LTS and derivatives macOS 12: Monterey macOS 13: Ventura (12 and above - macOS x86_64 Intel) and M1 (macOS arm64 - Apple silicon) (only the last two major versions are supported)	Yes	Free for everyone
SAS	Windows Windows Server Linux	Windows 10, 11 Windows Server - 2016, 2019, 2022	No	Yes (No limit)

TABLE 2 Supported Connections

Programs	Supported Connections
Alteryx	Adobe, Alteryx, Platforms, Amazon Platform, Anaplan, Apache, ASCII, Avro, Automation Anywhere, Box, CSV, Databricks Platforms, Denodo, Dropbox, dBase, ESRI Platforms, EXASOL, GIS, Excel(XLS/XLSX), Google Platforms, Gzip, HTML, IBM Platforms, JSON, MapInfo Platforms, Marketo, Microsoft Platforms, MongoDB, MySQL, NetSuite, OpenGIS, Oracle, Pivotal, PostgreSQL, PrestoDB, Qlik, SAP, SAS, ServiceNow, Pivotal Greenplum, PostgreSQL, Presto, Progress OpenEdge, Qubole Presto, Salesforce Platforms, Splunk Platforms, SAP Platforms, ServiceNow Platforms, SharePoint Lists Platforms, SingleStore, Snowflake, Spark SQL, Teradata Platforms, TIBCO Data Virtualization (Cisco Information Server), Vertica, Web Data Connector, Other Databases (JDBC), Other Databases (ODBC), Snowflake, SQLite, SRC Geography, Tableau, Teradata Platforms, Text, Vertica, UiPath, XML, Zip Files, Salesforce
Tableau PREP	Microsoft Platforms, Text Fiel, JSON, PDF, Spatial File, Tableau Platforms, Actian Vectorwise, Alibaba Platforms, Amazon Platforms, Anaplan, Apache Drill, Azure Platforms, Box, Cloudera Hadoop, Databricks, Denodo, Dremio by Dremio, Dropbox, Esri Connector, Exasol, Firebird 3, Google Platforms, Hortonworks Hadoop Hive, IBM Platforms, ImpalaIntuit Quick, Books Online, Kognitio, Kyvos, LinkedIn Sales Navigator, MariaDB, Marketo, Marketing Cloud Intelligence, Mark, Logic, MonetDB, MySQL, Odata, OneDrive, Oracle Platforms
RapidMiner	Oracle, Microsoft SQL Server, MySQL, PostgreSQL, Teradata, HP Vertica, IBM Netezza, MongoDB, Cassandra, Apache Solr, Splunk Dropbox, Amazon S3, Microsoft Azure Blob Storage, Salesforce, Twitter Mozenda, Zapier, MDB/ACCDB, XLS/XLSX XML, ARFF/XRFF, DBFSAV SAS DTA QVX
KINME	Amazon Redshift, Amazon Athena, Google BigQuery, H2, Microsoft Access, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, SQLite, Vertica, Apache Derby, Exasol, Google BigQuery, IBM DB2 / Informix SAP HANACSV, Excel, Text File, JSON, ZIP, Parquet
SAS	Databricks, Hadoop, IBM IAS, IBM Netezza Performance Server, MySQL, Microsoft SQL, Oracle, PostgreSQL, Teradata, Microsoft Azure Cloud Platforms, Google Cloud Platform, Amazon Web Services Cloud Platform, Oracle Cloud Platform, Microsoft Access, Microsoft Excel, HTML, Text File, Stata (*.DTA), SPSS (*.SAV), JMP

TABLE 3 Data, Statistical, and Language Capabilities

Programs	Data, Statistical, and Language Capabilities			
	Data Wrangling	Preliminary Statistics	ML Forecasting	Natural Language Processing
Alteryx	Yes	Yes	Yes	Yes
Tableau PREP	Yes	Yes	Not directly	Yes
RapidMiner	Yes	Yes	Yes	Yes
KINME	Yes	Yes	Yes	Yes
SAS	Yes	Yes	Yes	Yes

TABLE 4 Spatial and Reporting Capabilities

Programs	Spatial and Reporting Capabilities	
	Spatial/GIS	Reporting
Alteryx	Yes	Yes
Tableau PREP	Yes	No
RapidMiner	Yes	Yes
KINME	Yes	Yes
SAS	Yes	Yes

III. SUPPORTED CONNECTIONS, DATA WRANGLING, AND STATISTICS

All programs reviewed support common file types and connections such as CSV, XLSX, and MySQL or other database file types. Cloud connections, such as to Amazon Web Services, were also a ubiquitous feature. Several of the programs, SAS and Alteryx for example, have their own proprietary file types that they can connect to, while RapidMiner supports a range of file types but does not have a proprietary format of its own. Spatial or Geographic Information Services (GIS) data types were also common, with all programs capable of utilizing GIS data. Overall, Alteryx and Tableau support the widest range of file types and connections, while RapidMiner, KINME, and SAS were more limited in what the program can connect to.

Data wrangling was possible with all of the reviewed programs, though the ease of use between platforms in performing data wrangling was not evaluated as a component of this paper. As with Data wrangling, preliminary statistics could be performed with all programs except Tableau, but for SAS this step required writing code. In order to run preliminary statistics for Tableau a formula had to be created. For more advanced statistics, such as forecasting, Tableau was the only platform without forecasting capabilities. All platforms were, to some degree or with slight modifications, capable of natural language processing. In Tableau Prep, this is not directly available and users must employ the “Ask Data” function, while RapidMiner requires a text processing extension.

IV. FEATURE COMPARISON

To compare these software tools, we considered their capabilities in data wrangling, preliminary statistics, machine learning forecasting, open-source integration, license types, natural language processing, spatial GIS, and reporting

features. Although all five tools support these features, some present specific challenges.

Alteryx is a strong competitor for Windows users; however, its lack of support for macOS or Linux platforms is a significant disadvantage, as it restricts accessibility for users on these operating systems. Additionally, the one-year student license limitation, similar to Tableau's, can be challenging for students [6] [7]. Undergraduate programs typically span four years, while graduate programs often last at least two years, making the one-year license potentially insufficient for mastering the tool. On the positive side, Alteryx supports all features without requiring external extensions, simplifying its usage. Its user-friendly interface allows seamless management of features through built-in modules. Moreover, Alteryx provides excellent training resources, including well-prepared documentation and interactive tutorial videos, which enhance the user experience.

While Tableau Prep has a modern interface, certain features like machine learning forecasting and natural language processing require additional Tableau products. Furthermore, applying preliminary statistical methods in Tableau Prep often involves creating custom formulas instead of using predefined modules. Additionally, Tableau Prep does not directly support reporting functionality.

Both Alteryx and Tableau Prep offer a one-year license for students, but there are no time limitations for RapidMiner, KNIME, and SAS. RapidMiner provides an extensive range of features managed through extensions, which users may need to install as required. For Linux users, RapidMiner and SAS are particularly good options. [8] [9] [10]

KNIME offers a modern and visually appealing interface. While its commercial use requires a paid license, it is free for personal use without time limitations, making it a viable long-term option for students and individuals. KNIME integrates all tools internally, eliminating the need for external modules.

SAS, another strong contender for data analysts, has no time restrictions for student use, making it suitable for long-term adoption, similar to KNIME and RapidMiner. However, its reliance on a programming language instead of drag-and-drop modules can extend the learning process. Additionally, its older interface design has been subject to criticism, which may affect user satisfaction.

V. CONCLUSION

After reviewing five data analytics platforms, Alteryx, Tableau, KINME, RapidMiner, and SAS, university students have access to all were able to perform data wrangling, utilize GIS, and had natural language processing capabilities. Differences emerged in the range of file types or connections each program could use, with Alteryx and Tableau supporting a wider range of options than KINME, SAS, and RapidMiner. Additionally, for preliminary statistics or forecasting Tableau lacked functionality compared to the other platforms reviewed.

REFERENCES

- [1] Alteryx, "System Requirements," 2024. [Online]. Available: <https://help.alteryx.com/current/en/server/system-requirements.html#idp406292>.
- [2] Tableau, "Technical Specifications," 2024. [Online]. Available: <https://www.tableau.com/products/techspecs>.
- [3] RapidMiner, "RapidMiner Studio System Requirements," 2024. [Online]. Available: <https://docs.rapidminer.com/9.4/studio/installation/system-requirements.html>.
- [4] Knime, "KNIME Server Installation Guide," 2024. [Online]. Available: https://docs.knime.com/2021-06/server_installation_guide/index.html#_introduction.
- [5] SAS, "Hardware and Resource Requirements," 2024. [Online]. Available: https://documentation.sas.com/doc/en/itopscdc/v_058/itopssr/n0ampbltwqgkjknlj3qogztsbbu0.htm.
- [6] Alteryx, "Welcome to the Learners Page," 2024. [Online]. Available: <https://community.alteryx.com/t5/Learners/ghp/Students>.
- [7] Tableau, "Prep your data the Tableau way," 2024. [Online]. Available: <https://www.tableau.com/academic-prep#:~:text=Tableau%20Prep%20is%20available%20for,a%20new%20one%20year%20license..>
- [8] RapidMiner, "Rapidminer student licence activations step confusion," 2024. [Online]. Available: <https://community.altair.com/discussion/22780/rapidminer-student-licence-activations-step-confusion>.
- [9] Knime, 2024. [Online]. Available: <https://www.knime.com/knime-analytics-platform#:~:text=KNIME%20Analytics%20Platform%20is%20free,all%20popular%20machine%20learning%20libraries..>
- [10] SAS, "SAS OnDemand for Academics," 2024. [Online]. Available: https://www.sas.com/en_us/software/on-demand-for-academics.html#8239eb92-b4a3-4ef9-9306-f708a14164e9.

A Brief Look at Chatbots in Light of the Future of University Websites: Benefits, Opportunities, Applications, Challenges and Risks

Hakan Temiz
Computer Engineering
Faculty of Engineering
Artvin Çoruh University
Artvin, Türkiye
htemiz@artvin.edu.tr
0000-0002-1351-7565

Abstract— Chatbots are widely used on every platform and device that includes human-machine interaction. One of the potential application areas of chatbots is university websites. These frequently visited platforms provide plenty of information and services to countless people every day, including academic and administrative staff, active and prospective students. On the other hand, it is observed that these web platforms have not yet benefited sufficiently from the chatbot technologies. However, chatbots emerge as a much more dynamic, interactive, fast and easy tool compared to canonical web pages. These tools have the potential to accelerate and facilitate access to information; to perform frequent tasks effortlessly from a single point without getting lost in piles of pages; to be customizable for people and to provide a higher level of user interaction and experience.

This study begins by providing a brief review of chatbots and their historical development. Then, it continues by discussing the challenges experienced by university web platforms, the potential opportunities and benefits that chatbots can offer to these platforms, and the current technologies and challenges. It is clear that chatbots will be widely used and play critical roles in university web platforms in the near future.

I. INTRODUCTION

Chatbots offer a very fast, easy and dynamic interaction in textual, audio and visual forms. Their high-level capabilities have paved the way for the increasingly widespread use of chatbots on every platform that include human-machine interaction, from education to health, from finance to customer relationship management, and so on. University websites, as one of the most common and potential application areas of chatbots, have recently gained an attractive dynamism, speed, flexibility and high-level user experience. These sites, as frequently-visited platforms, provide a wide range of information and services to many people every day, including academic, administrative staff, active and prospective students. The richness and diversity of the features and functions offered by these sites cause them to be served in a dispersed and hierarchical structure. This distributed and relatively static structure makes it difficult to access information and causes cumbersome and inefficient use of functions and services. They also lack the dynamism of chatbots.

On the other hand, chatbots emerge as a much more dynamic, interactive, fast and easy tool, compared to ordinary web pages. These tools have the potential to accelerate and facilitate access to information; to perform frequent tasks

effortlessly from a single point without getting lost among piles of pages; to be customizable for people; and to provide a higher level of user interaction and experience. They also promote the quality of education and services, communication and interaction between the parties, increasing the satisfaction of all stakeholders. Besides, universities are likely to reach much wider audiences with these tools.

This study provides a concise review on the benefits, potential opportunities, application types, difficulties and risks that can be provided by chatbots to university websites. In this context, many aspects such as development environments and technologies, difficulties, risks and security measures are discussed in detail and the right design and solution approaches are presented. It also reveals how well-designed chatbots with the right content and scope can benefit all internal and external stakeholders and potential students. Additionally, the opportunities and potential that these tools can provide in improving the service quality of universities are discussed. Consequently, this study shows how chatbots will be an important part and play a critical role in the future of university websites.

II. RELATED WORKS

Thinking machines concept is based on Turing's article "Computing Machinery and Intelligence" [1]. The article evaluates the issue of whether machines can think like humans from various perspectives and also offers answers to possible opposing views. It proposes the "imitation game", which will later be called the "Turing Test". This game involves a questioning human trying to determine whether the person communicating via a textual interface is a human or a machine. If it cannot be determined whether the other person is a human or a machine, the machine is considered to be able to think. In this article, it can be said that the foundations of chatbots begin with the concept of machines [1].

The first chatbot ELIZA was developed by Weizenbaum in 1966 [2]. It is a rule-based system. chatbot communicates by reflecting the user's statements back in the form of a question, acting like a Rogerian psychotherapist. It faced problems such as producing illogical and irrelevant answers; not being able to develop conversations based on past conversations; establishing dialogue in the form of superficial and simple answers; not taking emotions into account, and the inadequacy of hardware and software technology of the period. Nevertheless, Weizenbaum, who saw that some people were trying to establish an emotional connection, was

disturbed by the emotional and ethical consequences that this tendency could create [3]. He emphasized that artificial intelligence lacked human values and emotional judgments and should not replace people in decision-making processes [4].

Elizabeth is an improvement of ELIZA in several aspects [5]. It implemented input rules, keyword patterns and various output rules. The main disadvantages of Elizabeth are that it does not have the ability to split and separate user input and to combine responses [6]. Colby et al. developed PARRY in 1971 for the purpose of modeling paranoid processes [7][8]. While ELIZA imitated a Rogerian therapist, PARRY attempted to verbally imitate a paranoid schizophrenic. A panel of judges tried to distinguish between real patients and the program by their textual communication, but observed that the program communicated very similarly to the patients [9]. It tried to give an "emotional response" by applying emotion weighting to the inputs. However, it lacked the ability to understand, adequately express emotions, respond quickly and learn [10]. PARRY is considered to be more advanced than ELIZA.

Racter, developed in 1983, could produce unique sentences and prose using context-independent grammar rules. Despite its lack of real understanding, its procedural text generation capabilities were a remarkable development for the time [11]. So much so that in 1984, he wrote "The Policeman's Beard is Half Constructed", which was touted as "the first book written by a computer program". It was remarkable that it produced expressions with internal meanings that appeared human-like. It was suggested that it could simulate human-specific emotional tones such as desires and dreams [11]. In the 1990s, Creative Technology developed Dr. Sbaitso (Sound Blaster Acting Intelligent Text to Speech Operator) for digitized voices. Dr. Sbaitso could convert text to speech, but with simple responses, he acted like a psychologist [12]. His inability to carry on a dialogue and his superficial responses prevented him from passing the Turing test.

In 1995, the first online chatbot, A.L.I.C.E. (Artificial Linguistic Internet Computer Entity), was developed using the AIML (Artificial Intelligence Markup Language) language. Improvements over the years have earned A.L.I.C.E. the Loebner award in the Turing Test competition in 2000, 2001 and 2004. While ELIZA had only 200 keywords and rules, A.L.I.C.E. had 41 thousand templates and related patterns [13]. However, it lacked deep understanding and could not express emotions in a human-like way. Albert One, created based on the ELIZA project, won the Loebner Award in 1998 and 1999 [10]. Albert One, which took a long time to develop, was developed using a huge database of conversational content, expressions and related responses.

In the 90s, one of the internet-based games, Tinymuds, allowed players to create and customize scenes, landscapes, etc. with a text interface, and software-controlled robots were called "bots". Its developer, Michael L. Mauldin, developed Chatterbot, which used the chat corpus in the game's pool to chat with players and generate answers to questions about paths, rooms, and objects [3]. Chatterbot performed pattern matching and gave random answers without repetition. Its most distinctive feature was that it imitated human behaviors and attitudes during conversation, such as misspellings, corrections, and late responses at variable speeds.

The first chatbot that could learn, Jabberwacky, was developed in 1988 with Cleverscript [14], aiming to provide a fun and interesting conversation simulation. It finds the most appropriate thing to say through contextual pattern matching techniques and improves its intelligence using the information it acquires. It was put on the Internet in 1997. Years later, a new version called Cleverbot [15] was able to have interesting and entertaining conversations. The ability to learn from conversations is an important development in showing that chatbots can be learned. This feature earned it the Loebner Award in 2005 and 2006.

In 2001, SmarterChild was developed, paving the way for future chatbots such as Apple Siri, MSN Messenger, Samsung S Voice. SmarterChild could interact via SMS and instant messaging (IM), AOL Instant Messenger (AIM), MSN and Yahoo Messenger, etc. It could provide data such as news, weather, stock prices, etc. In 2002, ZoeOnAOL was announced, which could communicate in Spanish and German and provide similar information. A similar chatbot was ZolaOnAOL, which aimed to provide humorous and entertaining chat. However, both were short-lived bots [16]. SimSimi, which tries to help people relieve their stress with interesting answers, was developed in 2002 [17]. Its distinctive feature is that it tries to learn by asking the user when it cannot understand the question or find the answer. On the other hand, the lack of a mechanism to distinguish whether the words are good or bad has caused controversy [18]. In 2003, the Learning and Organizing Cognitive Assistant CALO project was launched. The aim of the CALO project was to create a cognitive assistant that automatically learns and performs routine tasks from experience [11]. CALO inspired products such as Siri.

Suzette and Rosette, developed with ChatScript, a language that demonstrates stronger pattern matching and efficiency than AIML, won the Loebner Award in 2010 and 2011 respectively [19]. Their biggest challenge is that they require a tremendous amount of effort to produce hand-crafted pattern-based matching rules [20]. ChatScript's open source nature has provided support for natural language processing and conversation research. Rose won the Loebner award in 2014 and 2015 [20]. Mitsuku, which inherited the features of ALICE, was also developed with AIML [21] [11]. Its ability to show understanding, empathy and warm friendship was good for people who were lonely. In 2013 and between 2016 and 2019, it won the Loebner Award five times.

Siri, which has roots in the DARPA-funded Personalized Assistant that Learns (PAL) program, was developed by Apple for iOS in 2010 and is an important milestone [11]. Today, it is possible to use Siri on any iPhone, iPad, iPod, Apple Watch, Apple TV, and Mac with an internet connection [22]. It has many useful capabilities such as scheduling meetings, setting alarms, playing music, making phone calls, sending SMS, activating/deactivating phone functions, searching the web, and providing directions, by communicating in many languages. Despite having the ability to adapt and learn from the user's interactions as it interacts, it has not passed the Turing test.

IBM's supercomputer DeepQA managed to beat two Jeopardy players in 2011 [23]. It used natural language processing to analyze the question and extract key elements such as answer type and entity relationships. Machine learning (ML) algorithms were leveraged to analyze and process messy

and unstructured data. Although it failed the Turing test, it inspired subsequent work.

In 2012, Google developed Google Now, a virtual assistant that works with NLP and helps with things like weather, traffic alerts, reservations, financial information, sports scores, time, and location. Google Now was replaced by Google Assistant in 2016 [24].

Microsoft introduced the personal assistant Cortana in 2014 to help with tasks such as setting reminders, sending emails and messages, chatting, etc. with voice commands. However, it was discontinued due to various malicious flaws and was replaced by the AI-powered Copilot [25]. In 2014, Amazon launched the voice assistant Alexa. It offered a range of functions such as playing music, providing information such as weather, sports, news, etc., and controlling home devices through voice interaction with devices. Today, Alexa can serve on devices such as tablets, phones, televisions, etc. as a system hosted in the cloud. It serves as a smart home system consisting of a range of solutions that make life easier. Another bot, Panda Ichiro, developed also in 2014, has attracted people's attention with its entertaining style and has become quite popular. It can provide a series of practical information about daily life [26]. It has a series of remarkable skills such as telling jokes and some basic EQ skills. This feature later inspired the development of some components of Microsoft XiaoIce [27], which has the ability to establish long-term emotional bonds. XiaoIce can understand people's emotional states, moods, intentions, and their changes over time, and respond accordingly. Its ability to establish long-term emotional bonds differentiates it from recent chatbots such as Alexa, Google Assistant, and Cortana. While traditional chatbots focus on intellectual intelligence (IQ), XiaoIce exhibits emotional intelligence (EQ) and a personality/character in addition to IQ. In 2016, Microsoft introduced "Tay", which it claimed got smarter as it chatted with people. Within twenty-four hours of being released for testing, it began to learn people's racist, misogynist, and pro-Donald Trump tweets and produce similar responses [28], demonstrating the importance of a well-designed filtering mechanism.

In the last few years, there has been a serious acceleration in the development of content-generating artificial intelligence algorithms. These artificial intelligence algorithms are basically composed of ML models such as Generative Adversarial Networks (GANs) and transformers. These algorithms can generate a wide range of artificial content such as images, music, scenarios, poems, stories, and programming codes. For example, a well-trained can write code, summarize data, and create content such as images, music, videos, poems, etc. in line with the user's descriptions. With these capabilities, GPT (Generative Pre-trained Transformer) based chatbots, unlike typical chatbots, can process, understand, analyze different types of user data, and produce new content.

Replika, which is open source and the most popular and highly rated chatbot in Apple and Google Play stores, allows users to create customized avatars. It has prominent capabilities such as video calling and augmented reality experience. Its architecture is based on the GPT3 model, similar to Core Chat in XiaoIce [29]. Its human-like behavior and understanding enable people to form strong and long-lasting bonds. It is noted that forming such extreme emotional bonds can cause people to get hurt and even experience

serious psychological trauma. One of the most important of the mentioned generative networks is undoubtedly ChatGPT. ChatGPT, a Large Language Model (LLM) with a huge number of parameters and weights (175 billion in GPT-3), is capable of doing much more than traditional chatbots can do.

Google announced the virtual assistant Bard in early 2023 [30]. Its first version consisted of a lightweight model of Google's Language Model for Conversational Applications (LaMDA) technology. The built-in filtering mechanism contributed to its quality, reliability and robustness. In late 2023, it evolved into Gemini, with its new structure transformed into a multimodal LLM with language, voice, code and video understanding capabilities. Google Gemini is built on Pathways Language Model 2 (PaLM 2). Its neural network architecture is based on Transformer, the core component of recent chatbots such as ChatGPT. In addition to text and voice input, Gemini can directly understand and recognize images; it can parse and make inferences from complex visual content such as graphics, shapes, etc.

In the literature review, almost no chatbot developed for university automation was encountered. A notable study was conducted by Ülker et al. [31]. In this study, a chatbot application was developed that provides support for university administrators and academicians to perform various routine tasks and obtain information; for active students to access some services, obtain information about activities, perform tasks, receive help and obtain information; for candidate students to provide information about programs, staff, education-training activities, facilities, etc. Gaglio et al. developed a virtual assistant to provide information to students [32]. A similar attempt was made by Jangir and Ashok [33]. Another effort to develop chatbot for providing information was shown by Khan et al. It was aimed to answer and guide questions about the university through ML algorithms such as SVM and Random Forest [34]. Another application was developed by Shiwam et al. with WIT.AI to answer student queries [35]. [36] also tried to implement an informative chatbot for students. Chatbots other than these studies have been developed for some narrow scope interactions such as the university website and student information system rather than an integrated automation [37], [38].

III. CHATBOTS FOR UNIVERSITY WEB SITES

Higher education institutions have quite complex, distributed and large-scale data, information and processes. A large number of parties or stakeholders receive services, provide services and interact for more. Advanced software and hardware systems are used for the operation of the entire system. Web platforms also provide services to all parties, provide information, ensure that work is carried out, bring parties together, inform, advertise, introduce themselves, etc. and perform many other functions and services. These platforms are the front face, the showcase of these institutions. Web platforms consist of hundreds or even thousands of subpages in order to present a huge collection of data and information; manage work and processes; perform tasks, ensure coordination and perform other activities. Finding the desired one among the countless pages on these platforms, accessing information and performing automation are inherently difficult. On the other hand, chatbots and virtual assistants developed for the benefit of people to work on smart devices such as mobile phones, tablets, watches, IoT (internet of things) etc. and websites make daily life incredibly easier.

It is obvious that these software will also provide significant benefits on university web platforms.

Although there are successful big chatbots and virtual assistants around, such as Amazon Alexa, Apple Siri, Google Assistant, Microsoft Cortana, ChatGPT; they are not specialized in a particular application, such as our topic university websites. There is a significant need for chatbots specialized in certain areas. When we look at the literature, it is seen that there is very little effort to meet this need. The following sections discuss various aspects of the use of chatbots on university websites and also examine the actual situations of typical university websites.

A. Challenges Facing University Websites

University websites generally present the following challenges:

- **Single Tool:** Chatbots have the potential to integrate with all units, processes, systems, etc. of the university. Such a fully integrated system allows people to handle all their work and access information from a single point.
- **Natural and Multi-Channel Interaction:** Chatbots have the ability to interact through text, speech, dialogue components, visuals, etc. People can interact through any channel they want or need among these alternatives. This allows people to get their work done in the form of communication and interaction they use and are used to in daily life. An easier and more natural communication and interaction experience may well allow for the realization of a high-level university websites.
- **Repetitive Work and Productivity:** A significant portion of university activities consist of repetitive work and responding to questions (from both internal and external stakeholders) that reach a huge number. A significant number of employees (administrators, academics, etc.) and other resources are occupied with carrying out such work and answering frequently asked questions. This leads to a decrease in productivity. Chatbots can provide incredible convenience in performing repetitive work and answering frequently asked questions. It allows employees and other resources to be assigned to other tasks and increases productivity.
- **Speed and Accuracy:** Considering today's technology, traditional communication tools and university websites can sometimes be slow in accessing information and services. In addition, the complex and dispersed organizational structure of universities, consisting of multiple units spread over wide geographical areas, prevents the same up-to-date and accurate information from being available in all units. Therefore, incorrect and outdated versions of information are likely to circulate. On the other hand, accessing answers and services often becomes time-consuming and unbearable.
- **Access to Information and Services:** Although university websites contain the necessary services, processes and information on the pages, sometimes it is very difficult to access them. For example, although they are presented in logically grouped sets (menus, lists etc.), it is often seen that support has to be received by phone etc. because the information or

function presented on hundreds of pages cannot be found. Chatbots can easily and quickly provide the requested information or ensure that the activity is carried out through a single interface.

- **Access to the Outside World:** Universities use various communication tools (social media, email, instant messaging, etc.) and their own websites to introduce themselves to the outside world, announce their activities, and conduct marketing campaigns. These communication tools and environments are not always sufficient to achieve their goals.
- **Resource Limitation:** Regardless of its scale, every institution has limited resources such as personnel, information systems, etc. or has to use resources wisely. Limited resources cause disruption of work and services, delay in responses, decrease in service quality, dissatisfaction, etc.
- **Service Hours:** Although existence of university websites, which is the face of universities opening up to the world, is operating uninterruptedly, it is not possible to access every service, get information, and ask questions, etc. 24/7. Moreover, services and activities that require human power and intervention can only be carried out during working hours and to the extent that the current staff allows. However, today's "always open" life makes instant access difficult.
- **Communication and Interaction:** Traditional tools such as telephone, e-mail, etc. used by universities for the execution of services and activities and the transfer of information may be insufficient and cumbersome for the parties to establish effective and rapid interaction or may not serve the purpose. Third-party applications such as WhatsApp are frequently used for faster action and information transfer. In order to disseminate information, groups etc. are established or online meetings are tried to be reached to large audiences. On the other hand, chatbots have the potential to meet such needs.

Chatbots can be of significant benefit in overcoming the above-mentioned and other problems and deficiencies.

B. Benefits and Potential Opportunities That Chatbots Can Offer to Universities

The benefits and potential opportunities that chatbots can provide include, but are not limited to, the following:

- **Improving User Experience:** university websites serve a wide range of users, including prospective students, current students, parents, staff, alumni, and other organizations and people. These users frequently ask questions about campus life, the registration process, the academic calendar, departments, events, scholarships, grade information, administrative information, etc. Chatbots can improve the user experience by providing fast and accurate answers to FAQs. Quick access to requested information increases user satisfaction and prevents staff from being busy with repetitive tasks such as FAQs. Increased productivity.
- **Accessibility and 24/7 Service:** Traditional information and service processes are limited to office hours. Chatbots, on the other hand, can provide service to all parties, not just students, academics and

administrative staff, by working 24/7 without interruption. Accessibility is important for international people in different time zones. It can encourage applications, especially by allowing international student candidates to access more information about the university. In short, receiving uninterrupted service and information outside office hours and conducting business increases the satisfaction of all parties.

- **Richness in Interaction and Language:** The multi-language support of chatbots allows international students, academics and others to interact with universities quickly and effortlessly. There is no more convenient way or tool than chatbots for this purpose.
- **Contribution to Administrative Processes:** Most of the work in universities is repetitive and most of the requests for information are in the form of simple and repetitive questions. For example, routine questions and requests such as registration documents, application conditions, course programs, exam schedules and results, transcripts are among the jobs that frequently occupy administrative staff. Chatbots have the potential to automate administrative processes and take away especially repetitive and other time-consuming workloads of administrative staff. On the other hand, students, academicians and all other stakeholders also have the opportunity to receive fast support and reach results regardless of location and time.
- **Contribution to Interaction:** Chatbots have the potential to enable all stakeholders, such as student-student, student-academic, to interact with each other more easily, quickly, effectively and efficiently. In other words, they ensure that up-to-date information about current events, seminars, social activities, career opportunities, course programs, academic support, etc. activities and services reach all parties. In this way, they encourage all parties, especially students, to actively participate in university life. They help achieve a more active and dynamic university environment and life.
- **Personalization and Support for Learning:** Chatbots can support students not only with information such as course selection, academic calendar, exam dates and results, transcript information, etc. but also with their study processes, access to information, and learning styles. They can contribute to academic success by making personalized recommendations or offering study programs to individuals. They can also provide mentoring-like benefits to students by providing support and guidance for learning activities, processes, and career development. They can monitor students' educational experiences, help them see their weaknesses and strengths, and make suggestions.
- **Coordination:** Chatbots can enable all actors in the university (students, academicians, administrative staff, etc.) to coordinate with each other more quickly, easily, and powerfully. For example, organizing

meetings, sending emails to people, establishing groups to communicate collectively, assigning and collecting homework, etc. can be done easily.

C. Some Chatbot Initiatives for Universities and Educational Institutions

It is seen that almost all the software for developing chatbots for higher education institutions are produced by commercial entrepreneurs. Some of them are briefly mentioned below:

- Ivy.ai¹ allows creating chatbots for higher education institutions, in addition to other institutions.
- TARS² is a commercial chatbot developed to browse the university website, get information and answer questions.
- College Chatbot³ is a mini chatbot developed based on Python to answer questions with ML and NLP support.
- BoltBolt⁴ is a ChatGPT based chatbot developed for university registration and admission processes.
- Fynd Copilot⁵ is a ChatGPT based chatbot developed for university registration and admission processes.
- FastBots⁶ allows developing a chatbot that answers frequently asked questions to students and academics.

D. Risks and Challenges

- **Continuous Monitoring and Improvement:** In order for chatbots to work efficiently, they must be monitored regularly and updated in line with feedback. Feedback helps identify situations where chatbots have incomplete or incorrect understanding and answers. In addition, new questions and answers are added to the knowledge base of chatbots to improve the user experience. Continuous monitoring and improvement is the basis for chatbots to become more successful over time and to adapt to evolving and changing business processes and user interactions.
- **Data Security and Privacy:** Data security and privacy are among the most important priorities. It is essential that personal data is collected with a reliable and legally compliant infrastructure and technology, processed and protected in accordance with laws, standards and policies. Data integrity, leak-proofness etc. must be ensured by using technologies such as SSL, secure connection, hashing algorithms in communication. A system that does not include appropriate, sufficient and minimum security measures may cause loss of trust, various doubts and concerns, loss of reputation, time and money and various legal sanctions. It may cause withdrawal from using the chatbot system that will provide many potential benefits.
- **Difficulties in Staying Current:** The information provided by universities is frequently updated. For example, the academic calendar changes every semester, courses, course plans, course instructors, registration dates, academic and other staff change

¹ <https://ivy.ai/>

² <https://hellotars.com/blog/chatbot-for-college-website-handle-enquiries-seamlessly>

³ <https://github.com/roshancharlie/College-Chatbot-Using-ML-and-NLP>

⁴ <https://element451.com/blog/chatbot-for-admissions-and-enrollmentleverage-ai-for>

⁵ <https://www.copilot.live/usecase/chatbot-college-admissions>

⁶ <https://fastbots.ai/ai-chatbots-for-colleges>

constantly. Therefore, providing up-to-date status and data is not an easy task. In order to keep chatbot alive, this dynamic change must be constantly monitored and necessary actions must be taken. This requires additional resources, labor, etc.

- **Subject Expertise:** Although chatbots provide many benefits and make life incredibly easier, by their nature, they may not be a panacea. They cannot be expected to provide perfect service in complex and specific questions and situations. For example, there is no clear answer to which student should take which elective course. In such cases, the answers of chatbots alone may not be sufficient, and it is quite understandable that they may be inadequate.
- **Technical Support and Infrastructure:** Although the benefits of chatbots are countless, providing these benefits does not come for free. Designing, developing and commissioning these complex and advanced systems and continuing their life cycle is not as easy as it seems. It requires the coming together of experts. A certain amount of manpower and other resources must be allocated.
- **Malicious Attacks:** Let's imagine an integrated chatbot. Such a system should already take the necessary precautions for routine security and data privacy. In addition to these, all kinds of precautions should be taken against malicious or just curious attackers or actions that unintentionally damage a part or all of the system, infiltrate the system, seize data, etc. Otherwise, very serious consequences can be encountered regardless of the intention and the attacker. Taking countermeasures requires separate expertise, labor, time and money.

E. The Future of Chatbots

Chatbots today generally interact via text interface and additionally via voice. In some recent developments, chatbots that also take into account images and videos have begun to be encountered. It is almost certain that in the near future, we will establish a dialogue with smarter chatbots that interact via multiple channels including media, i.e. voice, image, text, video, virtual reality. In the future, there is no obstacle for the chatbots we use in other environments (at home, at work, etc.) to interact with university chatbots and perform useful activities for us.

IV. CONCLUSION

It is seen that chatbots, which have spread to every area of life and have produced amazing works, have only recently started to take place on university websites. However, chatbots have the potential to provide countless benefits on university web platforms. In the near future, chatbots will undoubtedly be the most important component of these websites.

In this study, in addition to the historical development of chatbots, the difficulties experienced on university web platforms, the potential opportunities and benefits that chatbots can provide, and the risks in the development and operation of chatbots are discussed. In addition, existing technologies are also examined. It is obvious that chatbots will undertake very important works on university web platforms in the near future.

REFERENCES

- [1] A. M. TURING, "I.—COMPUTING MACHINERY AND INTELLIGENCE," *Mind*, vol. LIX, no. 236, pp. 433–460, Oct. 1950, doi: 10.1093/mind/LIX.236.433.
- [2] J. Weizenbaum, "ELIZA—a computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 9, no. 1, pp. 36–45, 1966, doi: 10.1145/365153.365168.
- [3] M. L. Mauldin, "Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition," in *AAAI*, 1994, pp. 16–21. [Online]. Available: <https://cdn.aaai.org/AAAI/1994/AAAI94-003.pdf>
- [4] J. Weizenbaum, "Computer power and human reason: From judgment to calculation.," 1976.
- [5] A. Vishwakarma and A. Pandey, "A review & comparative analysis on various chatbots design," *Int. J. Comput. Sci. Mob. Comput*, vol. 10, pp. 72–78, 2021.
- [6] B. A. Shawar and E. Atwell, *A comparison between Alice and Elizabeth chatbot systems*. University of Leeds, School of Computing research report 2002.19, 2002.
- [7] K. M. Colby, S. Weber, and F. D. Hilf, "Artificial Paranoia," *Artif. Intell.*, vol. 2, no. 1, pp. 1–25, 1971, doi: [https://doi.org/10.1016/0004-3702\(71\)90002-6](https://doi.org/10.1016/0004-3702(71)90002-6).
- [8] K. M. Colby, F. D. Hilf, S. Weber, and H. C. Kraemer, "Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes," *Artif. Intell.*, vol. 3, pp. 199–221, 1972, doi: [https://doi.org/10.1016/0004-3702\(72\)90049-5](https://doi.org/10.1016/0004-3702(72)90049-5).
- [9] J. F. Heiser, K. M. Colby, W. S. Faight, and R. C. Parkison, "Can psychiatrists distinguish a computer simulation of paranoia from the real thing?: The limitations of turing-like tests as measures of the adequacy of simulations," *J. Psychiatr. Res.*, vol. 15, no. 3, pp. 149–162, 1979, doi: [https://doi.org/10.1016/0022-3956\(79\)90008-6](https://doi.org/10.1016/0022-3956(79)90008-6).
- [10] E. Adamopoulou and L. Moussiades, "Chatbots: History, technology, and applications," *Mach. Learn. with Appl.*, vol. 2, p. 100006, 2020, doi: <https://doi.org/10.1016/j.mlwa.2020.100006>.
- [11] M. Al-Amin *et al.*, "History of generative Artificial Intelligence (AI) chatbots: past, present, and future development," *arXiv Prepr. arXiv2402.05122*, 2024.
- [12] C. Owen, "Chatbots: The History, The Current Revolution, and The Future." [Online]. Available: <https://www.linkedin.com/pulse/chatbots-history-current-revolution-future-chris-owen/>
- [13] S. Kshirsagar, "A multilayer personality model," in *Proceedings of the 2nd international symposium on Smart graphics*, 2002, pp. 107–115. [Online]. Available: https://www.researchgate.net/profile/Nadia-Thalman/publication/228885591_A_Multilayer_Personality_Model/links/09e4151314d5d04ec400000/A-Multilayer-Personality-Model.pdf
- [14] R. Carpenter, "Jabberwacky." [Online]. Available: <http://www.jabberwacky.com/j2about>
- [15] R. Carpenter, "Cleverbot." [Online]. Available: <https://www.cleverbot.com/>
- [16] "AIM (software)." [Online]. Available: [https://en.wikipedia.org/wiki/AIM_\(software\)](https://en.wikipedia.org/wiki/AIM_(software))
- [17] S. Chandel, Y. Yuying, G. Yujie, A. Razaque, and G. Yang, "Chatbot: efficient and utility-based platform," in *Intelligent Computing: Proceedings of the 2018 Computing Conference, Volume 1*, Springer, 2019, pp. 109–122. [Online]. Available: https://www.researchgate.net/publication/328660686_Chatbot_Efficient_and_Utility-Based_Platform_Proceedings_of_the_2018_Computing_Conference_Volume_1
- [18] D. Sambandaraksa, "SimSimi chatbot banned in Thailand." [Online]. Available: <http://www.telecomasia.net/blog/content/simsimi-chatbot-banned-thailand>
- [19] M. L. McNeal and D. Newyear, "Introducing Chatbots in Libraries," *Libr. Technol. Rep.*, vol. 49, p. 5, 2013, [Online]. Available: <https://api.semanticscholar.org/CorpusID:60192623>
- [20] R. Jiang and R. E. Banchs, "Towards improving the performance of chat oriented dialogue system," in *2017 International Conference on Asian Language Processing (IALP)*, 2017, pp. 23–26. doi: 10.1109/IALP.2017.8300537.
- [21] "Mitsuku Chatbot: Unveiling the Award-Winning AI Companion." [Online]. Available: <https://fastbots.ai/blog/mitsuku-chatbot-unveiling-the-award-winning-ai-companion>

- [22] H. Soffar, "Apple Siri features, use, advantages, disadvantages and using of Siri For Learning." [Online]. Available: <https://www.online-sciences.com/technology/apple-siri-features-use-advantages-disadvantages-using-of-siri-for-learning/>
- [23] A. Lally and P. Fodor, "Natural language processing with prolog in the ibm watson system," *Assoc. Log. Program. Newsl.*, vol. 9, p. 2011, 2011.
- [24] S. Thakur, "Feed your need to know." [Online]. Available: <https://blog.google/products/search/feed-your-need-know/>
- [25] Pandasecurity, "Cortana security flaw means your PC may be compromised." [Online]. Available: <https://www.pandasecurity.com/en/mediacenter/mobile-news/cortana-security-flaw/>
- [26] N. OOSUGI, K. ITAZAWA, M. FUKUDA, S. SHIOZAWA, and M. IZUHARA, "A large scale chatterbot system implementation and its log data analysis," *JSAI Tech. Report, SIG-SLUD*, vol. 72, p. 13, 2014, doi: 10.11517/jsaislud.72.0_13.
- [27] L. Zhou, J. Gao, D. Li, and H.-Y. Shum, "The Design and Implementation of Xiaolce, an Empathetic Social Chatbot," *Comput. Linguist.*, vol. 46, no. 1, pp. 53–93, Mar. 2020, doi: 10.1162/coli_a_00368.
- [28] "Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day." [Online]. Available: <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>
- [29] I. Pentina, T. Hancock, and T. Xie, "Exploring relationship development with social chatbots: A mixed-method study of replika," *Comput. Human Behav.*, vol. 140, p. 107600, 2023, doi: <https://doi.org/10.1016/j.chb.2022.107600>.
- [30] S. Pichai, "An important next step on our AI journey." [Online]. Available: <https://blog.google/technology/ai/bard-google-ai-search-updates/>
- [31] Ü. Egemen, M. Ö. Bozbay, H. Ahioglu, and H. Temiz, "University Automation with Chatbot," in *1st International Multidisciplinary Student Congress in the Light of Science*, 2024, pp. 68–69. [Online]. Available: <https://ubsder.org.tr/wp-content/uploads/2024/08/OGRENCI-KONGRESI-KITABI.pdf%0A>
- [32] S. Gaglio, G. Lo Re, M. Morana, and C. Ruocco, "Smart Assistance for Students and People Living in a Campus," in *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2019, pp. 132–137. doi: 10.1109/SMARTCOMP.2019.00042.
- [33] L. K. Jangir, P. Ashok, and R. Bhatt, "MEERA Multifunctional Educational Expert in Real-time Assistance," 2020, [Online]. Available: http://www.ir.juit.ac.in:8080/jspui/bitstream/123456789/7080/1/MEE_RA_Multifunctional_Educational_Expert_in_Real-time_Assistance.pdf
- [34] Z. M. Khan, H. Rehman, M. Maqsood, and K. Mehmood, "Artificial intelligence based university chatbot using machine learning," *Pakistan J. Eng. Technol.*, vol. 4, no. 2, pp. 108–112, 2021.
- [35] K. Shivam, K. Saud, M. Sharma, S. Vashishth, and S. Patil, "Chatbot for college website," *Int J Comp Technol.*, vol. 5, no. 6, pp. 74–77, 2018, [Online]. Available: https://www.researchgate.net/publication/362546056_Chatbot_for_College_Website
- [36] F. E. da Cruz Maria, "MyCampus Chatbot-A Conversational Agent that offers information to IST Students," 2019, [Online]. Available: https://www-di.inf.puc-rio.br/~sardinha/Students/2019_81965_FranciscoMaria.pdf
- [37] O. C. Nwokonkwo, A. M. John-Otumu, and M. C. Onuoha, "Implementation of a Web-based Chatbot Assisted Services for Results Information System".
- [38] F. Hussain and M. A. Anwar, "Towards informed policy decisions: assessing student perceptions and intentions to use ChatGPT for academic performance in higher education," *J. Asian Public Policy*, pp. 1–28, 2024.

Düşük Hesaplama Kapasiteli Mikroişlemcilerde Blockchain Ağları Üzerinden Veri Paylaşımı ve Veri Gizliliğinin Geliştirilmesi

Seçkin CANBAZ
Bilgisayar Bilimi ve Mühendisliği
İstanbul Sabahattin Zaim Üniversitesi
İstanbul, Türkiye
seckin@canbaz.info

Gökhan ERDEMİR
Department of Engineering
Management and Technology
The University of Tennessee at
Chattanooga
Chattanooga, TN, USA
gokhan-erdemir@utc.edu

Tayfun ACARER
Deniz Ulaştırma ve İşletmeciliği Meslek
Yüksekokulu, İstanbul
Piri Reis Üniversitesi
İstanbul, Türkiye
tacarer@hotmail.com

Özetçe— Veri gizliliğinin ve güvenliğinin önemi IoT teknolojilerindeki ilerleme ve Endüstri 4.0'ın getirdiği yeniliklerin birçok alanda etkin olarak kullanılmaya başlamasıyla birlikte daha da artmıştır. Son dönemde hızla gelişen teknoloji, başta CPU ve GPU'lar olmak üzere yapay zeka destekli donanımların artması, en güçlü algoritmalar ile gizlenmiş olan verilerin bile eskiye göre daha çabuk ele geçirilebilir olmasına sebep olmaya başladı. Standart ve üzeri donanım kullanan cihazlarda bu işlemler kolayca yapılırken, düşük hesaplama kapasiteli mikroişlemcili mimarilerde ek güvenlik uygulamaları kullanmak sistem performanslarını oldukça düşürmektedir. Blokzincir teknolojilerinin kullanımı bu sorunun çözümü için önemli alternatiflerden bir tanesidir. Blokzincir ağlarının ihtiyaç duyduğu donanım mimarisi ve donanım yeterlilikleri, ağın desteklediği işletim sistemi mimarisi, enerji tüketimi, verimlilik gibi etkenler kullanılabilirliği ve performansı direkt etkileyen faktörlerdir. Bu çalışmada, farklı mikroişlemci mimarileri kullanan donanımlar üzerinde farklı blokzincir ağlarının performansı ve uygulanabilirliği incelenmiştir. Verinin blokzincir ağı üzerinden gönderilirken var olan donanım, işletim sistemi ve iletişim ağı güvenliğinin blokzincir ağı ile artırılmasını amaçlanmıştır.

Anahtar Kelimeler — blokzincir; veri paylaşımı; geliştirme kartı; IoT; raspberry pi; jetson; mini pc.

Abstract— The importance of data privacy and security has increased with the advancement of IoT technologies and the effective use of innovations brought by Industry 4.0 in many areas. Recently, rapidly developing technology, the increase in artificial intelligence-supported hardware, especially CPUs and GPUs, has started to cause even the data hidden with the most powerful algorithms to be captured more quickly than before. While these operations are easily performed on devices using standard and above hardware, using additional security applications in microprocessor architectures with low computational capacity significantly reduces system performance. The use of blockchain technologies is one of the essential alternatives for solving this problem. The hardware architecture and hardware capabilities required by blockchain networks, the operating system architecture supported by the network, energy consumption, and efficiency directly affect usability and performance. This study examined the performance and applicability of different blockchain networks on hardware using different microprocessor architectures. It is aimed to increase the security of existing hardware, operating systems, and communication networks with the blockchain network while sending data over the blockchain network.

Keywords — blockchain; data sharing; development board; IoT; raspberry pi; jetson; mini pc.

I. GİRİŞ

İnsanoğlu var oluşun başlangıcından bu yana, yaşamlarını kolaylaştırmak ve güvenliklerini sağlamak adına bir sürü alet tasarlamış ve kullanmıştır. Özellikle yakın dönemde 2. Sanayi Devrimi ile başlayan elektrikli aletlerin kullanımı ve bu aletlerin kullanımından edinilen faydanın fazla olması insanları bu alanda daha fazla çalışmaya ve daha fazla buluş yapmaya yöneltmiştir [1, 2]. 3. Sanayi Devrimi döneminde bilgisayarların ve bilgisayarlı aletlerin keşfedilmesi ve kullanımının yaygınlaşması bu dönemin Dijital Devrim olarak adlandırılmasına sebep olmuş ve bu alanda daha fazla çalışma yapılarak 4. Sanayi Devriminin çok hızlı bir şekilde gerçekleşmesine sebep olmuştur [1, 2, 3].

İçerisinde bulunduğumuz yüzyılda gerçekleşen birçok çalışma ile içerisinde bulunduğumuz 4. Sanayi Devrimi dönemi, teknolojinin gelişim hızının en yüksek değerine ulaştığı, yazılımsal ve donanımsal alanlarda çok fazla keşifin yapıldığı, insan-makine ortaklığının sonuçlarının hayatımıza pozitif olarak yansıdığı bir dönem olarak ilerlemektedir [4]. İnsan müdahalesinden bağımsız özerk çalışan elektrikli ürünler, blokzincir ağları, yapay zeka ve yapay zeka destekli yazılımlar bu dönemin son günlerdeki popüler örneklerinden birkaçı olarak verilebilir [4, 5, 6].

4. Sanayi Devrimi, Sanayi 4.0 ya da en bilinen adıyla Endüstri 4.0 döneminde cihazların özerk olarak çalışabilmeleri, birbirleri ile sürekli iletişim kurarak kendi kendilerini yönetebilmeleri gibi teknolojiler ortaya çıkmıştır [3, 7]. Bu teknolojiler sayesinde, üretimde insana olan ihtiyacın azalması dolayısıyla sürekli-kesintisiz üretimin ortaya çıkması gibi üretim anlamındaki verimli süreçler beraberinde çözülmesi gereken başka sorunları ortaya çıkartmıştır. Belirli bir görevi tamamlamak için kullanılan en az iki cihazın birbirleriyle iletişimi sırasındaki her bilgi veri olarak ele alınmaktadır. Bu veriler; cihaz bilgileri, parolalar, sensör verileri, üretim verileri gibi çok çeşitli bilgiler olabilir. Değer sırasına göre en düşükten en yükseğe kadar değer içeren her veri, kurulan sistemler için önem arz etmektedir. Verinin oluşma aşamasından bir görevin tamamlanması için kullanılmasına kadar olan zaman ve sonra da depolanma aşaması yönetilmesi gereken ayrı bir süreçtir. Veri gizliliği, verinin şifrelenmesi gibi süreçler geçmişten günümüze kadar önemliken Endüstri 4.0 ile birlikte artan veri miktarının çok olmasıyla birlikte daha fazla önem kazanmıştır [7, 8]. Anlamlı ya da anlamsız her türlü verinin analiziyle oluşan değerler sayesinde bir operasyon hakkında bilgi edinilebilir ya da o operasyonun geleceği hakkında tahminde bulunabilir. Dolayısıyla veri gizliliği ve veriyi saklayabilmek günümüz

teknolojisinde düzgün yönetilmesi gereken önemli bir süreç olarak karşımıza çıkmaktadır [8, 9].

Hızla gelişen yapay zekâ teknolojisi, günümüzde birçok özelliğiyle hayatımızı kolaylaştırıyor gibi gözükse de arkasında yatan teknoloji tamamıyla büyük veri setlerinden öğrenen ve öğrendiğini kullanan bir teknolojidir [5, 10]. Verinin en değerli maden olduğu dijital dünyada yapay zekâ destekli donanımlar veri güvenliği sağlamak ya da güvenlik altındaki verileri ortaya çıkarmak için kullanılmaya başlandı [11]. Merkezi işlem birimi-CPU ve grafik işlem birimi-GPU'ların dışında yeni bir teknoloji olan sinirsel işlem birimi-NPU'lar yapay zekâ araçlarının daha hızlı öğrenmesini ve dolayısıyla veri güvenliği sağlayan teknolojilerin daha kısa sürede savunmasız duruma gelmesini sağlamaktadır [12, 13].

Teknolojik alanda ürünlerin artmasıyla birlikte günümüzde enerji ihtiyacı çözülmesi zor olan büyük bir problem olarak hayatımızı tehdit etmektedir. Kullanılan elektrikli cihazların elektrik tüketim miktarının fazla olması, bizleri kaynakları kısıtlı olan Dünyamızda bulunan enerji kaynaklarının yanı sıra, yeni enerji kaynağı teknolojilerini araştırmaya itmektedir [14]. Bunun yanı sıra var olan cihazların enerji tüketimlerini düşürme yöntemi ise kısa zamanda daha hızlı ve etkili bir çözüm olarak öne çıkmaktadır [15].

Yapılacak işlemlere özel olarak üretilen elektronik cihazlar ile en az enerji tüketimi hedeflenmiş ve bu doğrultuda yeni teknolojiler ve cihazlar tasarlanmıştır. Günümüzde özerk olarak hareket edebilen araçlarda dahil olmak üzere yeni teknolojiler için yüksek kapasiteli bilgisayarlar yerine daha düşük hesaplama kapasiteli mimarileri kullanan bilgisayarlar tercih edilmekte ve cihazların enerji tüketimi en aza indirgenerek, kullanım zamanları ters orantılı olarak en fazla faydayı elde edecek şekilde artırılmıştır [15]. Bu tür bilgisayarların, işlem güçlerinin, depolamalarının ve enerji tüketimlerinin düşük olması bir takım güvenlik zafiyetlerine yol açmaktadır. Bunlardan en bilineni, en az iki cihaz arasındaki iletişimde, paylaşılan verinin, şifresiz olarak paylaşılması ya da zayıf şifreleme algoritmaları ile şifrelenerek paylaşılmasıdır [16]. Ayrıca cihazlar üzerinde güvenlik yazılımlarının kullanılmaması ya da zayıf yazılımların kullanılabilmesi verinin güçlü saldırılar karşısında savunmasız bırakılmasına sebep olmaktadır. Tüm bunlar paylaşılan verinin güvenliğinin yeterli şekilde sağlanamaması problemi [16, 17]. Bu sebeptendir ki, düşük hesaplama kapasiteli mimarilerde paylaşılan verinin güvenliğinin sağlanması çözümlenmesi gereken bir problemdir. İletişimde kullanılan ağın güvenliğinin sağlanması bir önlem olarak düşünülse bile iletişim ağına olabilecek sızıntılar veri güvenliği için tehdit oluşturabilir. Blokzincir teknolojisi kullanılarak paylaşılan verinin güvenliğinin artırılması bu tür tehditler için alternatif bir güvenlik önlemi olarak kullanılabilir [18]. Bu sayede verinin kendisi henüz iletişim ağına paylaşılmadan önce korunma altına alınarak verinin güvenliği daha üst seviyede sağlanabilmektedir.

Blokzincir ağları kurulurken en önemli konulardan biri üzerlerindeki verilerin dağıtık defter teknolojisi sebebiyle sürekli-değiştirilemez olması ve şifreli bir biçimde saklanarak güvene alınmasıdır [19, 20]. Hemen hemen her blokzincir ağı temelde aynı özellikleri taşımasına rağmen, kullandığı teknolojiler itibarıyla (programlama dili, defter teknolojisi, şifreleme algoritmaları vb.) belirli donanım kaynaklarına ihtiyaç duymaktadır. Bu donanım mimarilerinin güvenlik

seviyeleri, hesaplama kapasiteleri gibi kriterler ağı performansına direkt etki etmektedir. Dolayısıyla çalışmaya başlamadan önce iki türlü seçim yapılabilir. Bir yöntem, kullanılacak blokzincir ağına göre donanım seçimi iken diğer yöntem bunun tam tersi şekilde eldeki donanıma uygun bir blokzincir ağının seçilmesi olacaktır.

II. LİTERATÜR TARAMASI

A. Blokzincir Teknolojisinin Veri Güvenliğindeki Rolü

Blokzincir temelde bir veri tabanı mantığıyla çalışır. Klasik veri tabanından en temel farkı, içerisindeki veriler merkezi bir veri tabanında tutulmak yerine standart bir konumu olmayan, Dünya'nın herhangi bir yerinde bulunabilen gönüllü kullanıcıların cihazlarında depolanmaktadır. Kullanıcıların her biri eş olarak adlandırılır. Bu dağıtık yapı sayesinde bir veri değiştirilmek istenirse tüm eşlerdeki verilerin değiştirilmesi gerekir [21]. Bu eşlerin yüzbinlerce hatta milyonlarca olduğunu düşünürsek verilerin değiştirilmesinin imkansız yakın olduğunu söyleyebiliriz.

Blokzincir aslında blok ve zincir sözcüklerinin birleşiminden oluşan bir birleşik sözcüktür. Blokzincir bloklardan ve bu blokların birbirine bağlanması ile oluşan bir zincirden oluşur. Blokları temelde verilerin kaydedildiği veri kümeleri olarak adlandırabiliriz. Blokların sıralı şekilde art arda eklenmesiyle birlikte zincir yapısı oluşturulur [22]. Zincire yeni bir blok eklenilmek istediği zaman tüm eşlerin onayı alınmak zorundadır. Dolayısıyla zincire eklenen bir blok daha sonradan silinemez ya da değiştirilemez [21, 22, 23]. Onaylanmış eşler arasında bulunan zincirlerin herhangi birinde geçmişe yönelik yapılan bir değişiklik, diğer eşler tarafından kontrol edilir ve reddedilir. Bu işlem sonucunda bu eş artık geçersiz olacaktır.

Blokzincirin başlangıç bloğu Genesis Blok olarak tanımlanır. Bu başlangıç bloğu, projenin ilk oluşturulma aşamasında manuel olarak yazılır ve blokzincirin kurallarını belirleyerek temel yapısını oluşturur [24]. Genesis blok üzerine eklenecek her blok madenciler ile zincire eklenir ve bunlar belirli kurallara uymak zorundadır. Bir blokzincire blok eklenmesi süreci 3 ana adımda meydana gelir; blok oluşumu, işlemlerin doğrulanması ve var olan zincire eklenmesi [25].

Bir bloğun içerisinde bulunması gereken standart veriler vardır. Her blok iki ana bileşenden oluşur, başlık ve işlem kısmı. Başlık içerisinde birçok veri bulunmaktadır. İlk olarak bir önceki bloğun hash değeri bulunur. Her blok, kendisinden önce zincire kaydedilen bloğun hash değerini blok başlığının içerisinde belirterek sıralamayı belirler ve zincir içerisindeki bir blok üzerinde değişiklik yapılmasının önüne geçmiş olur [26]. Blok içerisindeki herhangi bir veride yapılacak tek bir karakterlik değişim bile bloğun hash değerinin tamamen farklı oluşmasına sebep olur. Bu da kendisinden sonraki bloğun başlığındaki o bloğa ait hash değerleri ile uyumsuzluk oluşturarak bloklar arası bağlantının kopmasına ve dolayısıyla zincirde bozulmaya sebep olur.

Başlık içerisindeki bir diğer veri zaman damgasıdır. Bloğun oluşturulduğu tarih ve saat bu zaman damgası içerisinde yer alır. Nonce değeri, madenciler tarafından kullanılır ve belirli bir zorluk seviyesine ulaşmak için oluşturulan rastgele bir sayıdır. Merkle Root, blok içerisindeki tüm işlemlerin kriptografik şekilde özetlenen biçimidir [26]. Başlıktan sonraki işlemler kısmı ise, blok içerisinde yer alan işlemlerin listesidir. İki blok oluşumu arasında geçen sürede

herhangi bir işlem gerçekleşmediyse, bloklar işlem verisi olmadan oluşturulur.

Blok oluştuktan sonraki adım onaylama adıdır. Farklı konsensus mekanizmaları için farklı onaylama yöntemleri vardır. Bloklar tasarlandıkları ağı konsensus mekanizmasına göre onaylandıktan sonra zincire eklenir. Zincire eklenen blok diğer tüm eşlere dağıtılır ve eşler bu bloğu onayladıktan sonra kendi zincirlerine eklerler. Bir sonraki adımda tüm eşler yeni zincir durumlarını veri tabanıyla senkronize ederek bir sonraki blok için hazır duruma geçerler [27, 28].

Yapılan tüm bu işlemler sayesinde, paylaşılacak istenen veri farklı güvenlik protokolleri uygulanarak gizlenir. Zincire yeni bloklar eklendikten sonra eski blokların değişmesi neredeyse imkansız hale geleceği için, zaman geçtikçe veriler üzerinde değişiklik yapılamaz. Dolayısıyla ağ üzerine paylaşılan veri güvenli, şeffaf ve değiştirilemez şekilde saklanır.

B. IoT ve Düşük Kaynaklı Cihazlarda Güvenlik Sorunları

Endüstri 4.0 ile birlikte hayatımızda çokça yer bulmaya başlayan bir terim var artık, sürdürülebilirlik. Bu terimin içerdiği değerlerin en önemlilerinden bir tanesi ise enerji tasarrufu ve enerjinin doğru kullanımı. Son zamanlarda teknolojik cihazların çok kullanılması, yapay zekâ araçlarının eğitilmesi, veri madenciliği yapılması, elektrikli araçların ortaya çıkması vb. gibi sebepler enerji tüketimini çok arttırmıştır. Enerjinin doğru kullanılması adına ise enerji tüketen cihazlar yeniden tasarlandı ve doğru enerji kullanımı ile sadece verilen görevi yerine getirebilecek kadar donanım kaynağına sahip cihazlar üretildi [29]. IoT cihazlar, mini bilgisayarlar, geliştirme kartları (Raspberry Pi, Jetson, vb.) bunlara örnek verilebilir. Düşük hesaplamasına kapasitesine sahip bu cihazlar, enerjiyi doğru şekilde kullanmaya başladılar fakat büyük bir sorunu beraberinde getirdiler, o da güvenlik [30].

Kullanılan donanım kaynaklarının kısıtlı bir işlem gücünü desteklemesi, güvenlik protokollerinin daha düşük seviyelerde çalışmasına, bazı durumlarda hiç uygulanmamasına ya da eski yöntemlerin uygulanmasına sebep olmaktadır. Günümüzde kullanılan modern ve güçlü şifrelemelerin bu cihazlar üzerinde yapılamaması kullanılan kriptografik protokollerin yetersiz şekilde uygulanmasına sebep oluyor. Standart dışı geliştirilen yazılımlar, açık kaynak kodlu işletim sistemleri üzerine güvensiz kaynaklardan kurulan yazılımlar ise güvenlik açıklarının oluşmasına sebep oluyor ve bu açıkları kapatacak güvenlik yazılımlarının olmaması ise cihazları savunmasız bırakıyor [31].

Raspberry Pi, Jetson gibi geliştirme kartları için özel olarak tasarlanan işletim sistemlerini kullanan kullanıcılar ise varsayılan parolaları çoğu zaman değiştirmedikleri için kaba kuvvet saldırılarında savunmasız halde kalabiliyorlar [32]. Ayrıca bu cihazlar için geliştirilen uygulamalar güvenlik testlerinden geçirilmeden dağıtılabilir ve sistem üzerinde ciddi açıklar bulunabilir. Cihazlar üzerinde Trusted Platform Module (TPM) veya Secure Enclave gibi donanıma bağlı güvenlik modüllerinin olmaması ve SSH-HTTP üzerinden gelen saldırılara açık olması başka bir güvenlik açığı olarak göze çarpmıştır [33]. Ayrıca bu şekilde cihazlara erişilmesi durumunda bu durum ağ içerisinde bulunan diğer cihazları da tehlikeye atmış oluyor.

C. Daha Önceki Çalışmaların İncelenmesi

Düşük hesaplama kapasitesine sahip cihazlarda veri paylaşımında veriler şifrelenerek paylaşılabilir. Geleneksel şifreleme algoritmaları yüksek işlem gücü gerektirdiğinden daha hafif şifreleme algoritmaları (Elliptic Curve Cryptography (ECC), ChaCha20, vb.) tercih edilebilir [34, 35, 36]. Bir diğer yöntem ise standart kullanılan erişim protokollerine ek güvenlik katmanların (TLS/SSL-DTLS) eklenmesi yöntemidir [36]. Veri paylaşımı esnasında kimlik doğrulama yöntemleri kullanılabilir. API anahtarları ya da OAuth2 tabanlı tokenler, çift faktörlü kimlik doğrulama (2FA) ya da her cihaza özel kimlik sertifikaları oluşturularak kimlik doğrulama sağlanabilir [37]. Bu şekilde veri paylaşımı sırasında araya ve ağa sızmalar engellenebilir.

Farklı bir yöntem ise cihazlar üzerinde minimum verinin saklanmasıdır. Cihazların aktarması gereken verileri aktardıktan sonra yok ederek, olası sızmalarda geçmişe dayalı verilerin erişilmesi engellenmiş olur. Ayrıca cihaz yazılımlarının güncel tutulması, bu cihazlar için sızma testlerinin yapılması ve işletim sistemi ya da uygulama üzerindeki açıkların kapatılması güvenliği arttıracaktır.

Cihazlar fiziksel müdahalelere karşı da savunulmalıdır. Açılıştaki zararlı uygulama yüklenmesi engelleyen secure boot ya da disk şifreleme için bitlocker ve benzeri araçlar ile cihazlara fiziksel müdahalede bulunarak verilerin ele geçirilmesi engellenebilir. Cihazlar üzerinde gömülü bellek kullanımı, donanımların epoksi ile kaplanması gibi çözümler fiziksel ataklara karşı caydırıcı çözümler olabilir.

III. METODOLOJİ

A. Blokzincir Ağlarının Seçimi

Düşük hesaplama kapasitesine sahip cihazlarda ağ seçimi yapılırken dikkat edilmesi gereken önemli kriterler vardır. Bunlardan en önemlileri;

1- *Blokzincir üzerindeki verilerin şifrelenmesi:* Blokzincir üzerinde tutulan veriler farklı algoritmalarla şifrelenebilir ya da açık halde bulunabilir. Kullanılacak ağın şifrelemesi ihtiyaca göre belirlenmelidir [38, 39].

2- *Blokzincir ağının kurulacak işletim sistemi ile uyumluluğu:* Blokzincir ağları tasarım aşamasında her işletim sistemini destekleyecek yapıda kurulmamış olabilir. Kimisi sadece Linux çekirdeğine sahip işletim sistemlerini desteklerken bir çoğu Windows, Mac OS, Linux tabanlı işletim sistemlerini destekleyebilir. Hatta bazı uygulamalar Android ve IOS işletim sistemine sahip cep telefonu, tablet gibi cihazlara da kurulabilir.

3- *Blokzincir ağının donanım gereksinimleri:* Blokzincir ağları tasarlanırken belirli işlemci mimarilerine göre tasarlanabilir. Intel ve AMD'nin kullandığı x86 ve x64 mimarileri ve ARM mimarisi en çok kullanılan işlemci mimarileri olduğu için genelde blokzincirler bunları destekleyecek şekilde tasarlanırlar.

4- *Blokzincir ağının depolama kapasitesi:* Blokzincir ağları tasarlanırken her bir bloğun tutacağı maksimum veri miktarı ile zincirinin maksimum boyutu belirlenir. Bazı blokzincir ağları kurulum aşamasında zincirlerin belirli bir geçmiş tarihe kadar doğrulanmış verilerini bilgisayara indirerek kurulum tarihinden itibaren veri doğrulamasına izin

verebilir. Dolayısıyla kurulum aşamasında büyük hafızalar gerekebilir. Standart bir blokzincirin ana ağı (mainnet) 100GB civarında bir veri tutar. Dolayısıyla ana ağ üzerinde çalışacaksanız depolama kapasitesi bir sorun olabilir. Test ağlarında ise durum farklıdır. Yerel bir test ağı kullanacaksanız düğüm kendi bilgisayarınızda başlayacağı için ilk kurulumda büyük bir depolama alanı gerekmez fakat bazı blokzincir ağları online test ağları kullanırlar ve yine belirli bir miktar veriyi test ağı başlangıcında bilgisayara indirmek gerekebilir.

5- *Blokzincir ağının enerji tüketimi:* Tercih edilen blokzincir ağı işlemleri gerçekleştirirken şifreleme, veri gönderme, onaylama ve depolama aşamasında ek kaynak kullanımına ihtiyaç duyduğu için enerji tüketimi fazla olabilir. Bu da farklı ağlar arasında tercih edilmeme sebebi olabilir.

Blokzincir ağları seçilirken kurulum dokümanları detaylı şekilde okunmalıdır. Maalesef ki birçoğunun bu başlıkları anlatan düzgün açıklanmış veri sayfaları yok ve ancak kurulum aşamasında ya da kurulumdan sonra bunları kendiniz test etmeniz gerekiyor. Çok tercih edilen ve çok kullanılan birkaç blokzincir ağı dışında detaylı açıklanmış doküman bulamayabilirsiniz. Bu ağlarda genelde düşük hesaplama kapasiteli mimariler için uygun olmuyor.

B. Cihaz Seçimi ve Donanım Özellikleri

Blokzincir ağları test edilirken düşük hesaplama kapasitesine sahip cihazlar kullanılmıştır. Bu cihazların özellikleri aşağıdaki tabloda verilmiştir. Tablo I'de bu sistemlerin teknik özellikleri gösterilmektedir.

TABLO I. BİLGİSAYARLARIN TEKNİK ÖZELLİKLERİ

Marka Model	Raspberry Pi 4	Raspberry Pi 5	Jetson Nano	Dell Wyse Dx0Q
CPU	Broadcom BCM2711 Quad-Core Arm Cortex A72 1.8GHz	Broadcom BCM2712 Quad-Core Arm Cortex A76 2.4GHz	Quad-Core Arm Cortex A57 MPCore 1.43 GHz	Quad core Amd GX 415GA 1.5GHz
GPU	Broadcom VideoCore VI	Broadcom VideoCore VII	NVIDIA Maxwell architecture with 128 NVIDIA CUDA® cores	AMD Radeon™ HD 8330E
RAM	2 GB	8 GB	4 GB	8GB
Depolama	64 GB Micro SD	64 GB Micro SD	64 GB Micro SD	500 GB Sata Disk

C. Veri Gönderimi İçin Blokzincir Tabanlı Yaklaşım

Cihazlar arasında paylaşılacak verilerin blokzincir ağı üzerinden gönderilmesi yöntemi, var olan ağ güvenliğine ek olarak, verinin kendisinin de güvenlik altına alınmasını ve ağ düzeyinde gerçekleşecek sızmalara karşı veriyi korumayı sağlar.

Bir blokzincir ağına veri gönderimi işlemi aşağıdaki adımlarla tamamlanır:

- 1- *Gönderilecek verilerin hazırlığı:* Cihazlardaki sensörlerden ya da çalışan uygulamalardan oluşturulan veri öncelikli olarak gönderilmek için hazırlanmalıdır. Metin şeklindeki veriler hızlıca dönüştürülebileceği gibi

ses ve video formatındaki veriler her blokzincir ağında oldukları gibi gönderilemeyebilir. Bu verilerinde metin biçimine çevrilmesi gerekebilir. Veri güvenliğini arttırmak için veri önceden şifrelenebilir, bu ek bir güvenlik katmanı sağlar. Büyük verilerin ise tamamı blokzincir üzerinde depolanmaz, bunun yerine verinin hash değeri oluşturularak ağ üzerinden bu hash değeri paylaşılır. Veriler ayrıştırılarak sadece gerekli verilerin gönderilmesi sağlanır, bu yöntem ile zincirde gereksiz verilerin tutulması engellenir.

- 2- *İşlem oluşturulması:* İşlemler, gönderici-alıcı adresleri ve işlem maliyeti gibi bilgileri içerecek şekilde düzenlenir [25, 40, 41, 42].
- 3- *İşlemin imzalanması:* Gönderici, veriyi kendi anahtarıyla imzalar. Bu sayede işlemin güvenliğini ve doğruluğunu sağlar [25, 40, 41, 42].
- 4- *İşlemin blokzincir ağına gönderimi:* Hazırlanan işlemler blokzincir ağına gönderilir ve işlem havuzunda alınarak bekletilir [25, 40, 41, 42].
- 5- *Doğrulama ve işlem ücretlerinin ödenmesi:* Ağ üzerinde bulunan doğrulayıcılar veya madenciler ile işlemler doğrulanır. Doğrulama tamamlandığında işlem ücreti tahsil edilerek bloğa dahil edilir [25, 40, 41, 42].
- 6- *İşlemin blokzincire kaydedilmesi:* Blok olarak eklenen işlemler daha sonrasında zincire dahil edilerek süreç tamamlanır [25, 40, 41, 42].

IV. DENEYSEL ÇALIŞMALAR VE BULGULAR

A. Deney Kurulumu

Çalışma sırasında kullanılan cihazlar Tablo I'de verilmişti. Bu çalışmada düşük hesaplama kapasiteli mimariler için blokzincir ağları seçilerek kullanılabilirlikleri incelenmiştir. Aşağıdaki tabloda kullanılabilirlikleri test edilecek ağlar belirleyici özellikleriyle birlikte verilmiştir.

B. Kullanılabilirlik Değerlendirmesi

Donanıma göre araştırmalar yapılırken blokzincir ağlarının kullanılabilirlik durumlarıyla ilgili yeterli ve detaylı bilgilere ulaşmakta zorluklar yaşadık. Birçok dokümanda kurulum adımları ve süreçleri genel hatlarıyla verilse de bu adımların genellikle güncel sistem ve yazılım gereksinimlerini tam anlamıyla yansıtmadığına şahit olduk. Özellikle Linux çekirdeğine sahip işletim sistemleri versiyonlarının sık sık güncellenmesi, blokzincir ağlarının kurulumunda kullanılan uygulamaların versiyonlarının da hızla değişmesine neden olmakta ve bu durum kurulum aşamasında beklenmedik hatalarla karşılaşmamıza yol açmaktadır. Daha da önemlisi, bu uygulamaların belirli bir versiyon desteğiyle çalışması gerektiği için, uyumsuzluk sorunlarını çözmek çoğu zaman çok fazla zaman almış veya bazı durumlarda tamamen çözümsüz kalmıştır.

Bunun yanı sıra, çoğu blokzincir ağının dokümanlarında gerek resmi web sitelerinde gerekse GitHub sayfalarında, donanım desteğiyle ilgili yeterince detaylı bilgiye yer verilmemesi ciddi bir problem teşkil etmektedir. Özellikle ARM tabanlı işlemciler için destek sunulmadığından, bu ağların ilk bakışta tüm sistemlere kurulabilir gibi görünmesine rağmen, gerçekte IoT cihazlar ile Raspberry Pi ve Jetson gibi geliştirme kartlarına kurulmasının mümkün olmadığını

gözlemledik. Bu durum, bu tür donanımlarla çalışan sistemlerin blokzincir ağlarını entegre etme süreçlerini büyük ölçüde kısıtlamakta ve araştırma süreçlerinde ciddi bir darboğaz oluşturmaktadır. Donanım ve yazılım arasındaki bu uyumsuzluklar hem akademik hem de endüstriyel araştırmalarda zaman kaybına ve kaynak israfına yol açmakta, blokzincir teknolojilerinin daha geniş bir yelpazede kullanılabilirliğini engellemektedir. Aşağıda, Tablo II'de bahsedilen blokzincir ağlarının kurulum adımları incelenmiş ve kullanılabilirlikleri hakkında bilgiler verilmiştir.

TABLO II. BLOKZİNCİR AĞLARININ TEKNİK ÖZELLİKLERİ

Ağ	Konsensüs Mekanizması	Akıllı Sözleşme	Geliştirme Dilleri	Enerji Verimliliği
IOTA	Tangle (DAG tabanlı, merkeziyetsiz ve madencisiz yapı)	IOTA Smart Contracts	Rust, Go, JavaScript, Python, C++	Çok yüksek (Madencisiz Tangle yapısı)
Multichain	Özel zincir, konsensüs ayarlanabilir (genellikle oylama ile)	Kısıtlı Akıllı Sözleşme Desteği	Python, C++, JavaScript	Yüksek (Özel ağ, düşük işlem yükü)
Chia	Proof of Space and Time (PoST)	Kısıtlı	Python (CLVM - Chialisp)	Orta (Disk alanına bağımlı, enerji tasarruflu)
Tezos	Liquid Proof of Stake (LPoS)	Var	Michelson, LIGO, SmartPy	Yüksek (LPoS ile verimli doğrulama)
Algorand	Pure Proof of Stake (PPoS)	Var	Python, Go, JavaScript	Çok yüksek (PPoS çok düşük enerji kullanır)
Komodo	Delayed Proof of Work (dPoW)	Var (Basit)	C, Python, JavaScript, Rust	Orta
Stellar	Stellar Consensus Protocol (SCP)	Var	Rust, Go, WASM Python, JavaScript	Yüksek (Hafif SCP algoritması)
EOS	Delegated Proof of Stake (DPoS)	Var	C++	Orta
Carda	Validity + Uniqueness	Kısıtlı	Java, Kotlin	Yüksek (Özel ağda düşük yük)
Geth	Eski versiyonlar Pow, güncel versiyonlar PoS	Var (Ethereum VM)	Solidity, Vyper, Go	Orta

IOTA: Çeşitli kurulum senaryoları olmasına rağmen en kolay kurulumu Docker üzerine Hornet ağının kurulumu şeklindedir [43]. Kendi web sitesindeki kurulum dokümanlarında önerilen donanım olarak en az 4 çekirdekli bir işlemci, 8 GB Ram ve SSD depolama donanımı olması gerekirken işlemci mimarisinden, minimum disk kapasitesinden bahsedilmemiştir ve tüm işletim sistemlerini desteklediğinden bahsetmektedir [44]. Kurulum adımlarını takip ederek blokzincir ağını kurduğumuzda, uygulama başarılı bir şekilde çalışmaktadır fakat iki cihaz arasında veri paylaşımı yapılamamaktadır. Test ağına bağlanmaya çalıştığımızda ise test ağı sunucularının aktif olmadığı testlerimizce anlaşılmıştır. Özetle, docker üzerinde uygulama

başlatılsa bile iki cihaz arası veri gönderiminde başarı sağlanamamıştır.

Multichain: Bitcoin'in altyapısı üzerine özelleştirilmiş bir kurulumu vardır. İzinli blokzincir uygulamalarını desteklemek için üretilmiştir. Community (ücretsiz) ve Enterprise (Ücretli) olmak üzere 2 sürümü vardır [45, 46, 47]. Donanım gereksinimi olarak en az 512 MB Ram, 1 GB depolama alanı gereklidir. Kurulum dokümanlarında tüm işletim sistemlerini desteklediği belirtilse de işlemci mimarisi için herhangi bir belirteç yoktur [45, 46, 47]. Yaptığımız testler sonucunda, Intel mimarisinde herhangi bir sorun yaşanmamış, bir cihazda blokzincir başlatılarak 2 farklı cihazdan blokzincire veri gönderim işlemi gerçekleştirilmiştir. Fakat Arm mimarisini desteklememektedir. Kurulum yapılamamaktadır.

Chia Network: Sürdürülebilirlik ve enerji verimliliğini ön planda tutan bir blokzincir ağıdır. Donanım özellikleri olarak en az 4GB Ram, harddisk üzerinde 400GB boş alanı olan bir laptop ya da masaüstü bilgisayar kurulum için yeterlidir [48]. Test ortamında kullanılan donanımlar en az donanım yeterliliklerini karşılamadığı için kurulum yapılamamaktadır.

Tezos: Literatür taramasında düşük enerji tüketimi sebebiyle incelemeye aldığımız Tezos ağı, kurulum aşamasında yüksek Ram ve Disk kapasitesine ihtiyaç duyması sebebiyle başarısız bir kurulum adımı gerçekleştirilmesine sebep olmuştur. En az donanım gereksinimleri; 2 işlemci çekirdeği, 16 GB Ram (8 GB dahili ve 8 GB swap kullanılabilir), 100 GB SSD depolama alanıdır [49]. Ram ve depolama alanı ihtiyacından dolayı donanımlarımız için uygun değildir.

Algorand: Enerji verimliliği ile ön plana çıkan algorand ağı güvenlik ve merkezi olmayan yapısıyla güvenliği desteklemektedir. Donanım özellikleri olarak en az 8 işlemci çekirdeği, 16 GB Ram ve 100 GB SSD gibi yüksek özelliklere ihtiyaç duyması sebebiyle kullandığımız donanım cihazları üzerine kurulamamıştır [50].

Komodo: Az enerji tüketen blokzincirler arasında yer alan Komodo ağına, donanım ihtiyacı sebebiyle kurulum aşamasında sorunlara yol açmaktadır. En az sistem gereksinimleri ihtiyacı karşılasa da ARM işlemci desteği olmadığı için geliştirme kartlarımıza kurulumu gerçekleştirilmemiştir [51].

Stellar: Yüksek enerji verimliliği ve hızlı veri transferi vaadinde bulunan stellar ağı merkezi olmayan yapısıyla da güvenliğini ön plana çıkarmaktadır. Bilgisayarlara kurulan test ağı ile online test ağına veri gönderip, herhangi bir konumdan tarayıcı yardımı ile verileri okuyabilirsiniz. Stellar test ağı donanım ihtiyacımızı karşılaması ve ücretsiz test imkanı sayesinde cihazlarımıza kurularak başarılı şekilde test ağına veri gönderimi gerçekleştirilmiştir [52].

EOS: Yüksek performans gerektiren, ölçeklenebilir olan fakat merkezi olmayan uygulamalar için tasarlanmış bir ağıdır. Düşük işlem maliyetlerinin olması ve işlem sürelerinin kısa olması dolayısıyla tercih edilmektedir. Cihazlarımıza kurulumu denenmiş fakat yüksek donanım gereksinimleri sebebiyle kurulamamıştır [53].

Carda: Yüksek enerji verimliliği içeren ayrıca gizlilik ve birlikte çalışabilirlik konularına odaklanarak tasarlanmış bir blokzincir ağıdır. Arm işlemci mimarisi desteği olmaması sebebiyle kurulum sırasında başarılı olunamamıştır [54].

Geth: Ethereum ağının istemcisi olarak çalışmaktadır. Bu ağa katılma, akıllı sözleşmeleri çalıştırabilme ve madencilik gibi imkanlar sunar. Tam düğüm çalışması için yüksek depolama kapasitesine ihtiyaç duyar (en az 650 GB) [55]. Test ağı bulunmaktadır ve deney ortamında test ağı cihazlarımıza başarıyla kurulmuştur.

C. Ağların Güvenlik Seviyelerinin İncelenmesi

Çalışma sırasında ağların cihazlarımıza kullanılabilirlik durumları incelenmiştir. Bu çalışmada düşük hesaplama kapasiteli mimariler kullanılmamasının sonucu olarak birçok ağ cihazlarımıza kurulamamış, dolayısıyla detaylı test edilememiştir. Fakat her bir blokzincir ağı araştırılırken detaylı incelemeleri yapılmış ve bahsedilen ağlarla ilgili detaylı bilgiler edinilmiştir. Bu bilgiler ışığında her bir ağın güvenlik seviyeleri incelenmiştir.

IOTA ağı Tangle adı verilen bir yapı kullanmaktadır. Tangle yapısı yüksek ölçeklenebilirlik sağlarken veri güvenliği kullanıcıların işlemleri doğrulaması ile sağlanır. Bu ağ üzerinde işlemler şifreleme ile güvence altına alınır ve veri gizliliği bu şekilde sağlanır. Tangle üzerindeki işlem geçmişini anonim değildir ve ek protokoller ile artırılabilir.

Multichain ağına katılım ve veri gönderimi iznildir, katılım ve veri gönderim izni olmayan eşler paylaşımında bulunamaz. Veri gönderim izni ve verileri okuma izinleri ayrıdır. Bir eş sadece veri gönderebilir ve tüm verileri okuyamaz. Dolayısıyla yetkilendirilmiş eşler verileri görebildiği için veri gizliliği ve güvenliği yüksek seviyede sağlanmış olur.

Chia ağında veri gizliliği için ek şifreleme mekanizmaları yoktur fakat blokzincir üzerindeki işlemlerde standart kriptografik yöntemlerin kullanılması veri güvenliği için yeterli olmaktadır.

Tezos ağı Zero Knowledge Proof (ZKP) teknolojisi sayesinde işlemler doğrulanırken veri detayları gizlenir. Bu sayede güvenliği ve gizliliği üst düzeye çıkarır.

Algorand ağında özel durumlarda izinli ağlar oluşturulabilir. Bu ağlar dışında standart şifreleme politikaları ile işlemler korunur.

Komodo ağı dPoW mekanizması sayesinde işlemlerini Bitcoin ağına bağlayarak ek güvenlik sağlar. Ayrıca Zero Knowledge Proof teknolojisini desteklediği için veri gizliliği üst seviyededir.

Stellar ağı işlem verilerini SHA-256 gibi standart bir şifreleme algoritması ile güvence altına alır. İşlemler şeffaftır ve blokzincir üzerinde görülebilir. Stellar gizlilik odaklı bir ağ olmadığı için işlemler kayıt altına alınırken herkese açık durumdadır. Ek güvenlik sağlamak için veriler stellar ağı yerine özel bir ağ oluşturularak orada saklanabilir. Ek gizlilik protokolleri olmadığı için çok hızlıdır ve işlem maliyetleri çok düşüktür.

EOS ağında işlemler kriptografik olarak imzalanarak yetkisiz kişilerin erişme ihtimallerine karşı korunmaktadır. Blokzincir ağı verilerin değişmezliğini ve bütünlüğünü garanti eder. İşlemler herkes tarafından görülebilir, gizlilik odaklı değildir. Bu sayede kısa zamanda çok fazla işlem gerçekleştirir ve hızlıdır.

Carda ağı izinli bir blokzincir ağı olduğu için güvenlidir. Ağa erişim ve ağ üzerinde yapılacak işlemler kontrol edilerek onaylanır. Veriler yalnızca ilgili taraflarca erişilebilir ve bu

sayede gizlilik artırılmış olur. Diğer ağ kullanıcıları verilere erişemez.

Geth (Ethereum) ağı Proof of Stake (PoS) mekanizmasına geçiş yaparak veri güvenliğini arttırmış ve enerji konusunda da verimlilik sağlamıştır. ZKP teknolojisini destekler. Özel ve anonim işlemlerin yapılabilmesi için altyapı desteği sağlar ve bu sayede güvenliğini arttırmış olur.

Özet olarak, Komodo, Tezos ve Geth ağları anonimlik ve veri gizliliği için çözümler sunar. Multichain ve Corda gibi ağlar izinli yapıları sayesinde kontrol ve kurumsal gizlilik gerektiren uygulamalarda kullanılabilir. Iota ve Stellar ise ölçeklenebilirliği ve şifreleme özelliklerini ön planda tuttuğu için IoT cihazlarda ve mikro ödeme sistemleri için verimli olabilir.

Blokzincir seçimi yapılırken kullanım amacına göre seçim yapılmalıdır. Güvenlik ve gizlilik kriterleri iyi analiz edilmeli ve uygulama tasarımı adımı ağların desteklediği yapılar incelenerek ihtiyaca göre karar verilmelidir.

V. TARTIŞMA

A. Avantajlar ve Dezavantajlar

Blokzincir ağı ile veri paylaşımı işlemi sistem performansına direkt etki eden bir işlemdir. Blokzincir ağının başlatılması, şifreleme adımları, blokzincir ağının özelliklerinin depolanması gibi işlemler donanım özelliklerine göre kaynakları yoğun şekilde kullanabilirler. İşlem oluşturma ve imzalama süreçlerinde işlemci gücüne ihtiyaç duyulması hem zaman olarak işlemlerin az da olsa gecikmesine, hem de enerji tüketiminin artmasına direkt etki eder. Ayrıca standart bir P2P işlemine göre ağ trafiğini biraz daha yoğun kullanabilir. Çoğu blokzincir ağının her işlem için işlem ücretlerinin olması da ek maliyet oluşturabilir.

Fakat tüm bunlara karşı önemli bir avantaj sağlamaktadır, o da güvenlik. Tüm bu çözümler, paylaşılan verinin güvenliğini sağlamak içindir ki günümüzde tüm sistemlerde ek güvenlik donanımları ve yazılımları kullanarak bu performans sorunları göz ardı edilebilmektedir.

B. Uygun Blokzincir Ağlarının Belirlenmesi

Her ne kadar kullanılan cihazlar düşük hesaplama kapasiteli mimarilere sahip bilgisayarlar olsa da blokzincir ağına göre donanımların yeniden oluşturulması, donanım ağının yeniden kurulması yüksek maliyetli olabilir. Dolayısıyla genellikle yapılan uygulamalar zaten var olan bir ortam üzerine tasarıldığı için, blokzincir ağları genellikle var olan cihazlara göre seçilmektedir. Bizde yaptığımız uygulamada halihazırda laboratuvarımızda var olan cihazlar üzerinde çalışacak uygun bir blokzincir araştırmasında bulduk. Heterojen bir yapımızın olması ise çok farklı donanımlara sahip ortamlarda çalışacak bir blokzincir ağı araştırmamıza olanak sağladı. Farklı donanım mimarilerinde bile stabil şekilde çalışan bir blokzincir ağı, ek donanım maliyeti oluşturmadan var olan sisteme entegre edilebilir ve var olan bir yapı için uygun maliyetli ek güvenlik hizmeti sağlayabilir. Çalışmamızda araştırılan blokzincir ağları, çalışmamızda kullanılan donanımlara verdikleri tepkilerle bu anlamda bir yol gösterici olarak, bu şekilde uygulama yapmak isteyenlerin vakit kayıplarını önleyebilir.

C. Uygulama Alanları ve Gelecekteki Çalışmalar

Düşük hesaplama kapasitesine sahip cihazlara blokzincir teknolojisinin entegrasyonu şeffaflık, verimlilik ve en

önemlisi güvenlik gibi alanlar açısından birçok özellik katmıştır. Akıllı evler-şehirlere, tarım ve sağlık sektörleri, endüstriyel otomasyon, finansal hizmetler gibi birçok sektöre kadar geniş bir uygulama alanı bulunmaktadır. Cihazlarda bulunan sensörler üzerinden alınan verilerin değiştirilemez şekilde saklanması, veri bütünlüğünü koruması, merkeziyetsiz bir yönetim şekli sunması güvenliği artırırken, akıllı sözleşmeler kullanılması sayesinde bu süreçler daha kaliteli şekilde yönetilebilmektedir.

VI. SONUÇLAR

Bu çalışma ile düşük hesaplama kapasiteli mimarilerde blokzincir ağı üzerinden veri gönderilerek veri gizliliğine ek bir güvenlik önlemi sağlanmış oldu. Sistem tasarlanırken dikkat edilmesi gereken noktaların başında donanım ve blokzincir uyumluluğu geliyor. Araştırılan birçok blokzincir ağı, düşük hesaplama kapasitesine sahip cihazlara kurulmuyor, kurulum yapılsa bile kaynak kullanımı çok yüksek olduğu için verimlilik konusunda sınıfta kalıyor. Uygun cihaz için uygun ağ seçimi ya da tam tersi belirli bir blokzincir ağı için uygun cihaz seçimi ile yola çıkmak doğru tercih olacaktır.

Veri gizliliği birçok şekilde sağlanabilir. Blokzincir bize bu noktada standart veri gizleme özelliklerinden biraz farklı bir yöntem sunmaktadır. Veriyi şifreli bir şekilde birçok eş ile paylaşarak ve bunu bloklar şeklinde bir zincirin halkası yaparak, verinin gelecekte değiştirilmesine engel olarak veri bütünlüğünü ve gizliliğini korumaktadır. Dolayısıyla verilerimiz gizli ve bir bütün şekilde ve en önemlisi değiştirilemediği için güvenilir bir kaynak olarak saklanmaktadır.

Blokzincir ve IoT'nin birlikte kullanılması ile gelecekte birçok işlem insandan bağımsız olabilir. Blokzincir altyapısının kullanılmasıyla birlikte; endüstriyel makine ve robotların kendi kaynaklarını tamamıyla yönetebilir hale geldiği bir ortamda, gelecekte ihtiyaç duyacakları hizmeti ya da kendi kaynaklarını satın alabilecekleri platformların olması gibi uygulamalar ya da elektrikli otonom araçların şarj istasyonlarında kendilerini şarj ettikten sonra ödemelerini bu yapı üzerinden yapması veya otonom şekilde sürdürülen bir e-ticaret sitesinde herhangi bir üründe stok azaldığında, güvenli bir şekilde üretici ile irtibata geçip ödemesini yapan ve stoklarını güncelleyen uygulamalarda iletişim ve ödeme alanlarında kullanılabilir.

VII. KAYNAKLAR

- [1] D. Günay, "Sanayi ve sanayi tarihi," Mimar ve Mühendis Dergisi, sayı. 31, sayfa. 8-14, Ocak 2002.
- [2] S. Berktaş, ve R. D. Oraklıbel, "Sanayi Devrimi İle Gelen Değişim: İş Bölümü Ve Yabancılaşma," Atlas Sosyal Bilimler Dergisi, sayı. 6, sayfa. 112-121, Şubat 2021.
- [3] N. S. Pamuk ve M. Soysal, "Yeni sanayi devrimi endüstri 4.0 üzerine bir inceleme," Verimlilik Dergisi, sayı. 1, sayfa. 41-66, 2018.
- [4] O. Z. Fırat ve S. Ü. Fırat, "Endüstri 4.0 yolculuğunda trendler ve robotlar," İstanbul Üniversitesi İşletme Fakültesi Dergisi, cilt. 46, sayı. 2, sayfa. 211-223, 2017.
- [5] Z. Jan, F. Ahamed, W. Mayer, N. Patel, G. Grossmann, M. Stumptner ve A. Kuusk, "Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities," Expert Systems with Applications, cilt. 216, makale no. 119456, 2023.
- [6] G. Bovenzi, G. Aceto, V. Persico ve A. Pescapé, "Blockchain Performance in Industry 4.0: Drivers, use cases, and future directions," Journal of Industrial Information Integration, makale no. 100513, 2023.

- [7] S. A. Göv ve D. Erdoğan, "Dördüncü Endüstri Devriminin (Endüstri 4.0) Neresindeyiz?," İstanbul Gelişim Üniversitesi Sosyal Bilimler Dergisi, cilt. 7, sayı. 2, sayfa. 299-318, 2020.
- [8] T. P. Raptis, A. Passarella ve M. Conti, "Data management in industry 4.0: State of the art and open challenges," IEEE Access, cilt. 7, sayfa. 97052-97093, 2019.
- [9] A. Albugmi, M. O. Alassafi, R. Walters ve G. Wills, "Data security in cloud computing," 2016 Fifth International Conference on Future Generation Communication Technologies (FGCT), sayfa. 55-59, Ağustos 2016.
- [10] N. Polyzotis, M. Zinkevich, S. Roy, E. Breck ve S. Whang, "Data validation for machine learning," Proceedings of Machine Learning and Systems, cilt. 1, sayfa. 334-347, 2019.
- [11] T. Tan ve G. Cao, "Efficient execution of deep neural networks on mobile devices with NPUs," 20th International Conference on Information Processing in Sensor Networks (Co-located with CPS-IoT Week 2021), sayfa. 283-298, Mayıs 2021.
- [12] K. J. Raval, N. K. Jadav, T. Rathod, S. Tanwar, V. Vimal ve N. Yamsani, "A survey on safeguarding critical infrastructures: Attacks, AI security, and future directions," International Journal of Critical Infrastructure Protection, cilt. 44, makale no. 100647, 2024.
- [13] E. Bertino, M. Kantarcioglu, C. G. Akcora, S. Samtani, S. Mittal ve M. Gupta, "AI for Security and Security for AI," Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy, sayfa. 333-334, Nisan 2021.
- [14] A. Uzun ve M. Değirmen, "Endüstriyel işletmelerde enerji verimliliği ve enerji yönetimi," Uluslararası Ekonomik Araştırmalar Dergisi, cilt. 4, sayı. 2, sayfa. 83-97, 2018.
- [15] A. Abusukhon, B. Hawashin ve M. Lafi, "An efficient system for reducing the power consumption in offices using the Internet of Things," International Journal of Advances in Soft Computing & Its Applications, cilt. 13, sayı. 1, 2021.
- [16] A. Hameed ve A. Alomary, "Security issues in IoT: A survey," 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), sayfa. 1-5, Eylül 2019.
- [17] M. A. Razzaq, S. H. Gill, M. A. Qureshi ve S. Ullah, "Security issues in the Internet of Things (IoT): A comprehensive study," International Journal of Advanced Computer Science and Applications, cilt. 8, sayı. 6, sayfa. 383, 2017.
- [18] D. Minoli ve B. Occhiogrosso, "Blockchain mechanisms for IoT security," Internet of Things, cilt. 1, sayfa. 1-13, 2018.
- [19] M. A. Khan ve K. Salah, "IoT security: Review, blockchain solutions, and open challenges," Future Generation Computer Systems, cilt. 82, sayfa. 395-411, 2018.
- [20] A. Dorri, S. S. Kanhere, R. Jurdak ve P. Gauravaram, "Blockchain for IoT security and privacy: The case study of a smart home," 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), sayfa. 618-623, Mart 2017.
- [21] S. S. Sarmah, "Understanding blockchain technology," Computer Science and Engineering, cilt. 8, sayı. 2, sayfa. 23-29, 2018.
- [22] A. S. Ber, "Blokzincir (Blockchain) Teknolojisi Kapsamında Elektronik Çek," Journal of Marine and Engineering Technology, cilt. 2, sayı. 1, sayfa. 1-20, 2022.
- [23] A. S. Gaikwad, "Overview of blockchain," International Journal for Research in Applied Science and Engineering Technology, cilt. 8, sayı. 6, sayfa. 2268-2270, 2020.
- [24] E. E. Şahin ve O. Özkan, "Asimetrik volatilitenin tahmini: Kripto para Bitcoin uygulaması," Bilecik Şeyh Edebali Üniversitesi Sosyal Bilimler Dergisi, cilt. 3, sayı. 2, sayfa. 240-247, 2018.
- [25] H. Gül, "Blokzincir (Blockchain) Teknolojisi ve muhasebe bilgi sistemine etkileri," Sosyal, Beşeri ve İdari Bilimlerde Yenilikçi Yaklaşımlar, sayfa. 186-195, 2019.
- [26] M. Tanrıverdi, M. Uysal ve M. T. Üstündağ, "Blokzinciri teknoloji nedir? Ne değildir?: Alanyazın incelemesi," Bilişim Teknolojileri Dergisi, cilt. 12, sayı. 3, sayfa. 203-217, 2019.
- [27] A. F. Mendi, "Blokzincir mimarisine ve getirdiği fırsatlar," Avrupa Bilim ve Teknoloji Dergisi, sayı. 29, sayfa. 181-186, 2021.
- [28] M. Aslan ve M. C. Kasapbaşı, "Blok Zinciri Platformları, Fikir Birliği Mekanizmaları ve Ağın Güvenlik Analizi," Haliç Üniversitesi Fen Bilimleri Dergisi, cilt. 5, sayı. 1, sayfa. 43-72, 2022.

- [29] F. Akın ve S. Dinçer, "Endüstri 4.0 ve Çevresel Sürdürülebilirlik Arasındaki İlişki: Avrupa Birliği Ülkeleri Üzerine Ampirik Bir Analiz," *Yönetim Bilimleri Dergisi*, cilt. 22, sayı. 54, sayfa. 2449-2469, 2024.
- [30] İ. Avcı, "Akıllı evlerde IoT teknolojileri ve siber güvenlik," *Avrupa Bilim ve Teknoloji Dergisi*, sayı. 34, sayfa. 226-233, 2022.
- [31] C. Atac ve S. Akleyek, "A survey on security threats and solutions in the age of IoT," *Avrupa Bilim ve Teknoloji Dergisi*, sayı. 15, sayfa. 36-42, 2019.
- [32] İtigiç, "Raspberry Pi varsayılan şifreyi kaldırarak maksimum güvenlik sağlıyor," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://itigic.com/tr/raspberry-pi-removes-the-default-password-for-maximum-security/>
- [33] Secromix, "SSH ve Pentest: Güvenli Kabuğun Güvenlik Açıkları," [Çevrim İçi]. Erişim Tarihi: Dec. 15, 2024. Erişim Adresi: <https://secromix.com/blog/ssh-ve-pentest-guvenli-kabugun-guvenlik-aciklari/>
- [34] F. B. Pak, "Lale: Yeni bir hafif sıklet blok şifreleme algoritması tasarımı," *Yüksek Lisans Tezi*, Sakarya Üniversitesi, 2023.
- [35] Ö. Yel, "Akıllı ev cihazlarının haberleşmesinde hafif sıklet şifreleme algoritmalarının performans analizi," *Yüksek Lisans Tezi*, Sakarya Üniversitesi, 2024.
- [36] S. Böge, "Sanal özel ağlarda veri güvenliği," *Data Security on Virtual Private Networks*, 2018.
- [37] F. Fernández, A. Alonso, L. Marco ve J. Salvachúa, "A model to enable application-scoped access control as a service for IoT using OAuth 2.0," *2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN)*, sayfa. 322-324, Mart 2017.
- [38] A. Baliga, "Understanding blockchain consensus models," *Persistent*, cilt. 4, sayı. 1, sayfa. 14, 2017.
- [39] C. Saraf ve S. Sabadra, "Blockchain platforms: A compendium," *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, sayfa. 1-6, Mayıs 2018.
- [40] TÜBİTAK BİLGEM, "Blokzincir Nedir?," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://blokzincir.bilgem.tubitak.gov.tr/blokzincir-nedir>
- [41] G. Ünal ve Ç. Uluç, "Blok zinciri teknolojisi," *Bilişim Teknolojileri Dergisi*, cilt. 13, sayı. 2, sayfa. 167-175, 2020.
- [42] G. F. Türker ve K. Tanyeri, "Blokzincir Teknolojisi ile Nesnelerin İnterneti Tabanlı (IoT) Sistemlerin Veri Güvenliğinin Sağlanması," *Gazi University Journal of Science Part C: Design and Technology*, cilt. 1, sayı. 1, 2024.
- [43] IOTA, "Hornet: Welcome," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://wiki.iota.org/hornet/welcome/>
- [44] IOTA, "Hornet: Getting Started," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: https://wiki.iota.org/hornet/getting_started
- [45] Multichain, "Enterprise Features," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://www.multichain.com/enterprise/>
- [46] Multichain, "Community Edition Download," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://www.multichain.com/download-community/>
- [47] S. Ismail, H. Reza, H. K. Zadeh ve F. Vasefi, "A blockchain-based IoT security solution using multichain," *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, sayfa. 1105-1111, Mart 2023.
- [48] Chia, "Farming Guide: Obtain Hardware," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://docs.chia.net/farming-guide/#obtain-hardware>
- [49] OpenTezos, "Deploy a Node: Installation Guide," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://opentezos.com/node-baking/deploy-a-node/installation/>
- [50] Algorand, "Run a Node: Setup and Installation," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://developer.algorand.org/docs/run-a-node/setup/install/>
- [51] Komodo Platform, "Documentation: Komodo," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://komodoplatform.com/en/docs/komodo/>
- [52] Stellar, "Data Requirements Documentation," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://developers.stellar.org/docs/data/requirements>
- [53] EOS Network, "Node Operation: Hardware Requirements," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://docs.eosnetwork.com/docs/latest/node-operation/getting-started/hardware-requirements>
- [54] R3 Corda, "Host Prerequisites for Node Setup," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://docs.r3.com/en/platform/corda/4.8/enterprise/node/setup/host-prereq.html>
- [55] Ethereum, "Getting Started: Hardware Requirements," [Çevrim İçi]. Erişim Tarihi: 15 Aralık, 2024. Erişim Adresi: <https://geth.ethereum.org/docs/getting-started/hardware-requirements>

Plant Disease Recognition Using Transfer Learning: A Comparative Study of CNN Models

Özkan Nedzip
Electrical-Electronics Engineering
Vienna University of Technology
Vienna, Austria
oe.nedzip@gmail.com

Erdal Alimovski
Computer Engineering Departments
Istanbul Sabahattin Zaim University
Istanbul, Türkiye
erdal.alimovski@izu.edu.tr

Abstract—Plant disease detection and classification are critical for ensuring sustainable agriculture and food security. In recent years, convolutional neural networks (CNNs) have emerged as state-of-the-art solutions for image classification tasks due to their remarkable performance in extracting meaningful features from complex datasets. This study evaluates the performance of state-of-the-art CNN architectures, including VGG16, InceptionV3, and MobileNetV2, on a publicly available plant disease dataset. The dataset consists of high-resolution images of leaves affected by various diseases, as well as healthy samples, providing a robust testbed for assessing model performance. The results indicate that MobileNetV2 achieves the best overall performance, with the highest precision (94%), recall (94%), F1-score (94%), and test accuracy (94.21%). These findings establish MobileNetV2 as the most robust and efficient model for plant disease classification.

Keywords—Plant disease, recognition, cnn, agriculture.

I. INTRODUCTION

The precise and timely identification of plant diseases plays a vital role in ensuring sustainable agriculture and avoiding unnecessary expenditure of financial and other resources. In some cases, plant diseases may not exhibit visible symptoms, making it essential to utilize advanced analysis methods for detection and diagnosis [1]. The majority of plant diseases manifest visible symptoms, and the conventional method for diagnosis involves experienced plant pathologists identifying the disease through careful visual examination of affected plant leaves [2]. Accurate diagnosis of plant diseases by a plant pathologist requires strong observational skills to recognize characteristic symptoms. However, the vast diversity of plant species, variations in disease progression caused by climate change, and the rapid spread of diseases to previously unaffected regions can make it challenging for even experienced plant pathologists to identify certain diseases accurately [3]. The availability of expert and intelligent systems capable of accurately diagnosing plant diseases offers significant benefits to agronomists. Additionally, providing such a system enables non-expert farmers, especially those lacking access to agronomic and phytopathological support, to effectively identify and manage plant diseases [4].

Advances in artificial intelligence technologies have enabled the development of automated systems capable of delivering faster and more accurate results in disease diagnosis. Today, AI-based systems are widely used for the automatic detection of various diseases [5]. Over the past decade, numerous traditional machine learning models have been proposed for the detection and classification of plant diseases. For instance, in [6] authors employed Support

Vector Machines (SVM) for the early diagnosis and classification of diseases in sugar beet based on spectral plant indices. In addition to traditional machine learning methods, deep learning approaches have gained significant attention due to their superior performance in image classification tasks, making them particularly effective for detecting and classifying plant leaf diseases. Therefore, the importance of applying deep learning methods increases, as they offer enhanced accuracy and efficiency in handling complex tasks such as image classification, particularly in areas like plant disease recognition [7, 8].

In this study, three widely used Convolutional Neural Network (CNN) models, namely VGG16, ResNet50, and MobileNetV2, were applied to a publicly available plant disease dataset. The performance of these models was evaluated and compared in terms of accuracy, precision, and recall to determine their effectiveness in classifying plant diseases. The structure of the paper is as follows: Section 2 presents the utilized dataset and the methods, Section 3 details the experiments conducted, and Section 4 concludes the study.

II. MATERIALS AND METHODS

A. Dataset

The publicly available PlantVillage dataset covers totally 54,035 images. It includes **15 classes**, covering various plant species and their corresponding diseases, along with healthy leaf samples. This dataset serves as a valuable resource for training and evaluating image classification models. The original image resolution is 256x256 pixels. Some samples of the PlantVillage dataset are shown in Figure 1.

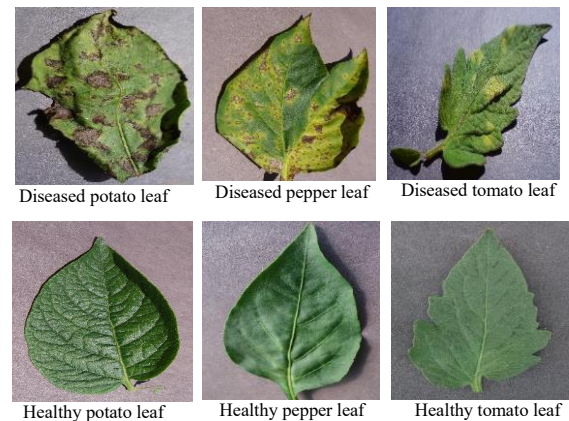


Fig. 1. PlantVillage Dataset Samples.

B. VGG16

The VGG [9] network is a deep convolutional neural network architecture, developed by researchers from the Visual Geometry Group (VGG) at the University of Oxford and Google Deepmind and was ranked among the top 5 models in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [10]. It is characterized by its depth, consisting of 16 layers with a uniform structure of small 3 x 3 convolutional kernels and 2 x 2 pooling layers, including 13 convolutional layers and 3 fully connected layers. The architecture has several versions with different numbers of layers. The two most common current VGGNet versions are VGGNet-16 and VGGNet-19. In this study, the classic VGG16 network serves as the foundation. Its simple yet effective design allows for efficient feature extraction, making it suitable for a wide range of image classification tasks. Additionally, its uniform architecture provides a strong baseline for comparison with other deep learning models. The VGG16 model's ability to learn hierarchical features makes it particularly valuable for identifying intricate patterns in plant disease images. Furthermore, its consistent performance and well-documented structure make it a reliable choice for benchmarking and model evaluation in deep learning studies.

C. InceptionV3

The InceptionV3 [11] model is a deep CNN architecture designed to be computationally efficient while maintaining high accuracy and performance on image classification tasks. The architecture is built of convolutional layers, pooling kernels and inception modules. The inception module is a block that applies multiple types of filter size to comprehend a variety of features at different scales and resolutions. The modular design achieves superior efficiency and scalability, this makes it a preferred choice for environments with limited computational resources. In this study, InceptionV3 was utilized to leverage its efficiency and multi-scale feature extraction capabilities for plant disease detection. The model's ability to process high-dimensional image data while minimizing computational overhead makes it particularly effective for identifying subtle disease patterns. Additionally, its proven performance across various image classification tasks ensures reliable and accurate results, even for challenging plant disease datasets.

D. MobileNetV2

The MobileNetV2 [12] model is a deep Convolutional Neural Network (CNN) architecture designed for high efficiency and performance, particularly in image classification tasks. It is optimized for mobile and edge devices with limited computational resources. The architecture incorporates depthwise separable convolutions, which significantly reduce computational cost by factorizing standard convolutions into depthwise and pointwise operations. Additionally, MobileNetV2 introduces inverted residual blocks with linear bottlenecks, enabling efficient feature extraction while maintaining accuracy. This lightweight and modular design makes MobileNetV2 a preferred choice for resource-constrained environments. Its balance of computational efficiency and accuracy allows it to perform well even on low-power devices. In this study, MobileNetV2 was employed to take advantage of its lightweight architecture for fast and reliable plant disease detection. Furthermore, its ability to deliver robust performance with fewer parameters makes it ideal for large-scale deployment. The model's adaptability ensures it can be

effectively integrated into mobile applications, providing practical solutions for real-time plant disease identification for plant pathologists and non-expert farmers.

III. EXPERIMENT

The experiments were conducted using Python 3.6 on a Linux operating system. TensorFlow [13] and OpenCV [14] libraries were employed for the implementation of each model. All experiments were carried out on a high-performance gaming laptop equipped with an 11th Gen Intel® Core™ i7-11800H CPU (2.30 GHz), 16 GB of RAM, and an NVIDIA GeForce GTX 1650 GPU with 4 GB of memory.

For the training and evaluation process, the dataset was divided into three subsets to ensure a reliable assessment of the models' performance. Specifically, 80% of the data was allocated for training, which was used to optimize the model's parameters and learn the underlying features of the images. 10% of the data was set aside for validation, allowing us to monitor the model's performance during training and fine-tune the hyperparameters to prevent overfitting. The remaining 10% of the data was reserved for testing, providing an independent evaluation of the model's ability to generalize to unseen data.

Each model was evaluated based on its performance metrics, including accuracy, precision, recall, and F1 score, to ensure a comprehensive assessment of their effectiveness in classifying plant diseases. The equations for the evaluation metrics are as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1 Score} = 2x \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

The performance of the three deep learning models—VGG16, InceptionV3, and MobileNetV2—was evaluated based on training accuracy, training loss, and test accuracy. The results indicate that MobileNetV2 outperformed the other models, achieving the highest training accuracy of 84.98% and the lowest training loss of 35.35%, leading to a test accuracy of 94.21%. In comparison, VGG16 attained a training accuracy of 83.45% with a moderate training loss of 37.44%, resulting in a test accuracy of 89.00%. On the other hand, InceptionV3 exhibited the lowest training accuracy at 71.19% and the highest training loss at 40.87%, yielding a test accuracy of 87.37%. These results highlight the efficiency and superior generalization capability of MobileNetV2, making it particularly suitable for resource-constrained environments where performance and accuracy are critical.

TABLE I. Performance Comparison of each model based on Accuracy, Loss, and Test Results.

Methods	Training Accuracy (%)	Train Loss (%)	Test (%)
VGG16	83.45	37.44	89.00
InceptionV3	71.19	40.87	87.37
MobileNetV2	84.98	35.35	94.21

The performance of the models was further evaluated in Table II. using precision, recall, and F1-score metrics to ensure a comprehensive analysis. The results demonstrate that MobileNetV2 achieved the highest scores across all three metrics, with a precision, recall, and F1-score of 94%. In comparison, VGG16 recorded a precision of 89%, a recall of 87%, and an F1-score of 87%, indicating consistent performance. Meanwhile, InceptionV3 showed slightly lower values, achieving a precision of 89%, a recall of 82%, and an F1-score of 84%. These results highlight the superior performance of MobileNetV2 in accurately detecting and classifying plant diseases, making it the most effective model among the three evaluated.

TABLE II. Performance Comparison of each model based on Precision, Recall and F1-score.

Methods	Precision (%)	Recall (%)	F1-score (%)
VGG16	89	87	87
InceptionV3	89	82	84
MobileNetV2	94	94	94

The Figure 2. illustrates the confusion matrix for the classification of plant diseases, showcasing the model's performance across various categories. Each row in the matrix represents the actual class, while each column represents the predicted class. The diagonal elements highlight the number of correctly classified instances for each class, where a higher value indicates better performance. Off-diagonal values represent misclassifications, showing where the model incorrectly predicted a different class. From the figure, it can be observed that the model performs well on most classes, as evidenced by the strong diagonal presence, particularly for categories such as *YellowLeaf_Curl_Virus* (325), *Tomato_Bacterial_spot* (197), and *Tomato_healthy* (149). However, slight misclassifications are present in certain classes like *Tomato_Early_blight* and *Potato_Early_blight*, where the model occasionally confuses them with other similar disease categories. The confusion between visually similar classes suggests a need for further refinement, such as incorporating additional features or improving image quality.

The results in Figure 3. show that the InceptionV3 model performs well across most classes, achieving high correct classifications for categories such as *YellowLeaf_Curl_Virus* (325), *Tomato_healthy* (153), and *Tomato_Bacterial_spot* (185). These classes exhibit minimal misclassifications, demonstrating the model's effectiveness in capturing distinct features and patterns.

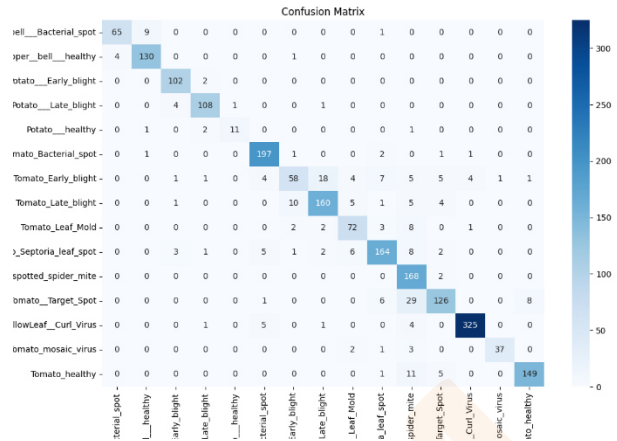


Fig. 2. Confusion Matrix of VGG16 model.

However, certain misclassifications are observed, particularly in visually similar classes. For instance, *Tomato_Early_blight* has 7 samples misclassified as *Tomato_Late_blight*, and *Potato_Early_blight* shows confusion with *Potato_Late_blight*. Similarly, *Tomato_Leaf_Mold* and *Late_blight* display slight overlaps, indicating challenges in distinguishing subtle variations in symptoms. These results suggest that, while InceptionV3 delivers strong overall performance, further optimization or additional data augmentation techniques may be required to reduce misclassification rates for closely related categories.

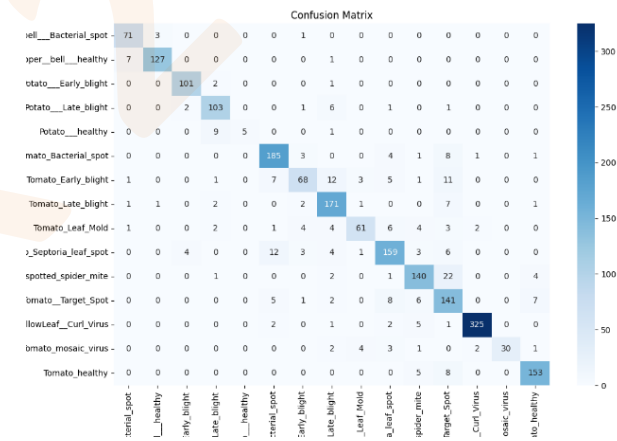


Fig. 3. Confusion Matrix of InceptionV3 model.

The results in Figure 4. indicates that the MobileNetV2 model delivers strong performance across most classes, with high correct classifications in *YellowLeaf_Curl_Virus* (333), *Tomato_healthy* (161), and *Tomato_Bacterial_spot* (194). These results demonstrate the model's capability to effectively distinguish these classes with minimal errors, reflecting its strong feature extraction and classification abilities.

However, some minor misclassifications are observed in classes with visual similarities. For instance, *Tomato_Early_blight* has 6 samples misclassified as *Tomato_Late_blight*, and *Potato_Early_blight* shows slight confusion with *Potato_Late_blight* (8 samples). Similarly, *Tomato_Leaf_Mold* and *Target_Spot* exhibit a small number of overlaps. These misclassifications suggest that further enhancements, such as improved data augmentation or fine-tuning, may help refine the model's performance in distinguishing closely related categories.

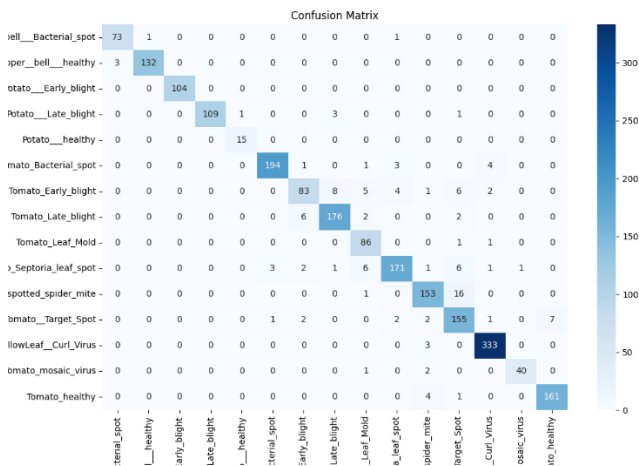


Fig. 4. Confusion Matrix of MobileNetV2 model.

The overall evaluation of the models VGG16, InceptionV3, and MobileNetV2 reveals that MobileNetV2 performs the best in distinguishing plant diseases with high precision and consistency. MobileNetV2 effectively handles feature extraction across different disease classes, minimizing errors and achieving reliable results. VGG16 also delivers strong performance but shows slight difficulty in differentiating visually similar diseases. On the other hand, InceptionV3 demonstrates relatively lower performance, particularly in classes with overlapping features, indicating its challenges in capturing finer details compared to the other models. Overall, MobileNetV2 stands out as the most effective model due to its robust classification capability and efficient handling of diverse plant disease categories.

IV. CONCLUSION

The evaluation of VGG16, InceptionV3, and MobileNetV2 highlights MobileNetV2 as the most effective model for plant disease detection and classification. MobileNetV2 excels in precision and consistency, efficiently extracting features across diverse disease classes while minimizing errors. VGG16 also demonstrates strong performance but struggles with visually similar diseases, whereas InceptionV3 exhibits lower accuracy, particularly in cases with overlapping features, indicating its limitations in capturing finer details. These findings emphasize MobileNetV2's robustness and efficiency, making it a promising solution for practical applications in plant disease detection. Future work can focus on further enhancing these models to address subtle disease patterns and improve overall classification performance.

REFERENCES

[1] Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, 61, 101182.

[2] Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and electronics in agriculture*, 72(1), 1-13.

[3] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf

image classification. *Computational intelligence and neuroscience*, 2016(1), 3289801.

[4] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, 145, 311-318.

[5] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4).

[6] Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and electronics in agriculture*, 74(1), 91-99.

[7] Barbedo, J. G. (2018). Factors influencing the use of deep learning for plant disease recognition. *Biosystems engineering*, 172, 84-91.

[8] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. *IEEE Access*, 9, 56683-56698.

[9] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[10] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115, 211-252.

[11] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

[12] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).

[13] TensorFlow Developers. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems* [Software]. Google. <https://www.tensorflow.org/>

[14] Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. <https://opencv.org/>

Performance Evaluation of Traditional and Deep Learning-Based Face Detection Algorithms

Merve Özer

Computer Engineering Department
Istanbul Sabahattin Zaim University
Istanbul, Türkiye
030122063@std.izu.edu.tr

Erdal Alimovski

Computer Engineering Department
Istanbul Sabahattin Zaim University
Istanbul, Türkiye
erdal.alimovski@izu.edu.tr

Asiye Demirtaş

Electrical-Electronics Engineering
Department
Istanbul Sabahattin Zaim University
Istanbul, Türkiye
asiye.demirtas@izu.edu.tr

Abstract— Face detection is a crucial task in computer vision, with applications ranging from security to human-computer interaction. In this study, we evaluate and compare the performance of three widely used methods: Haar Cascade, Histogram of Oriented Gradients (HOG), and Multi-task Cascaded Convolutional Neural Network (MTCNN) on the LFW dataset. The evaluation is based on accuracy and processing time. Experimental results show that MTCNN achieves the highest accuracy of 99.99%, failing to detect only one image, while Haar Cascade and HOG exhibit lower accuracies of 99.57% and 99.27%, respectively. Processing time analysis reveals Haar Cascade as the fastest, with an average detection time of 13.45 ms per sample, while HOG and MTCNN take longer, at 66.12 ms and 59.62 ms, respectively. These results highlight MTCNN's robustness in accuracy, making it suitable for applications requiring high precision, though with slightly higher processing time.

Keywords—Face Detection, Haar Cascade, Histogram of Oriented Gradients, Multi-task Cascaded Convolutional Networks.

I. INTRODUCTION

Face detection plays an important role in the field of computer vision and is the basis for many applications such as security systems [1], biometric verification [2] and human-computer interaction [3].

By detecting human faces in images or video streams, face detection facilitates advanced tasks such as face recognition [4], emotion recognition [5], and age or gender classification [6]. These capabilities are essential in fields such as security and surveillance, personalized user interactions, healthcare diagnostics, and entertainment technologies, where precise face analysis improves both functionality and outcomes. Despite its critical role, face detection faces numerous challenges, particularly in natural or uncontrolled environments. Variability in lighting, complex backgrounds, occlusions (such as glasses, hats, or masks), and diverse facial expressions can significantly impact detection accuracy. Additionally, differences in skin tones, facial orientations, and image resolutions present further obstacles. Overcoming these challenges requires robust algorithms capable of generalizing across diverse scenarios while maintaining computational efficiency, especially in real-time applications.

Face detection methods can be broadly categorized into two types: traditional approaches, which depend on manually designed features and classifiers, and deep learning approaches, which leverage advanced learning-based algorithms for improved accuracy and adaptability.

Conventional face detection methods were based on extracting features from images and applying classifiers to identify face regions. Techniques such as Haar cascade [7] and Histogram of Oriented Gradients (HOG) [8] - Support Vector Machine (SVM) [9] were significant milestones in classical face detection [10]. However, these methods face limitations in accuracy when dealing with complex variations in challenging datasets.

Modern face detectors significantly outperform traditional cascaded classifiers, achieving better results with advanced methods like Faster R-CNN [11], YOLO [12], and Single Shot Detector (SSD) [13]. Additionally, detectors such as Multitask Cascaded Convolutional Networks (MTCNN) [14] have demonstrated remarkable success, offering state-of-the-art performance on datasets like WIDER FACE.

In this paper, we performed face detection using traditional approaches, such as Haar Cascade and HOG, as well as deep learning approaches like MTCNN. In addition, we analyzed the performance of these methods on the LFW dataset.

The paper is organized as follows: Section 2 describes the dataset and methods utilized, Section 3 details the experiments performed, and Section 4 concludes the study.

II. MATERIALS AND METHODS

A. Dataset for evaluation

The Labelled Faces in the Wild (LFW) [15] dataset consists of a total of 13,233 facial images belonging to 5,749 individuals. The images were taken from the web in natural settings, showing variations in pose, lighting, and age. The majority of the images in the LFW dataset are RGB color images with a resolution of 250 x 250 pixels, with only a few exceptions.



Fig. 1. Example of Labeled Faces in the Wild Dataset

The images are collected from natural, real-world settings, introducing variations in pose, lighting, expressions, occlusions, and age. This diversity closely mirrors real-life scenarios, making the dataset an excellent benchmark for testing the robustness and accuracy of face detection, recognition, and verification algorithms.

B. Haar Cascade

Haar Cascade [16] is a machine learning-based object detection method widely used for tasks such as face detection [17]. It relies on Haar-like features, which are simple rectangular features that calculate the difference in pixel intensities between adjacent regions, capturing essential structural characteristics of objects. Figure 2 demonstrates examples of Haar-like features, where rectangular regions are used to extract intensity contrasts, such as horizontal or vertical edges and line structures. These features are applied to various positions and scales on the input image, capturing the distinct patterns associated with the object of interest. The method leverages an integral image to enable fast computation of these features and employs the AdaBoost[18] algorithm to select the most relevant ones for classification. A cascade of classifiers is then constructed to progressively eliminate non-object regions while focusing computational effort on promising areas.

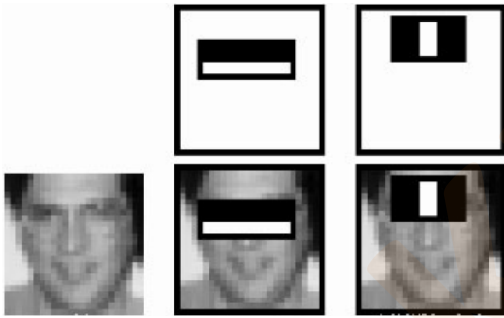


Fig.2. Illustration of features selected by AdaBoost, highlighting differences in intensity between the eyes, cheeks, and nose bridge.

C. Histogram of Oriented Gradients

Histogram of Oriented Gradients (HOG) is a feature descriptor used in image processing and computer vision, especially for object recognition. A feature descriptor extracts important information from an image and ignores unnecessary details to create a simpler representation of the image. This method was first introduced by Robert K. McCannell in 1986, but became widely used after being optimized for human detection by Dalal and Triggs in 2005 [8].

The computation of the Histogram of Oriented Gradients (HOG) descriptor involves a systematic process to extract gradient-based features for image analysis. This process can be described in the following stages:

Data Preparation: The first step in computing HOG is making sure the images are the right size. Images are usually

resized in a 1:2 ratio. For example, a 250x250 pixel image is resized to make it easier to process. The image is then divided into smaller cells of 8x8 or 16x16 pixels. These cells are used to compute the gradient information.

Grayscale Image and Gradient Calculation: HOG typically works with grayscale images (black-and-white images). By ignoring the color information and using only the brightness values, more accurate results can be obtained. The gradients in the image are calculated based on changes in pixel values, both in size and direction. The combination of the images magnitude and angle produces the gradient. The gradient is calculated as following [19]:

$$G = \sqrt{G_x^2 + G_y^2} \quad (1)$$

Where G_x and G_y represent the gradients along the horizontal and vertical directions, respectively. The gradient direction θ is determined using the following equation:

$$\theta = \arctan\left(\frac{G_x}{G_y}\right) \quad (2)$$

Creating Histograms and Dividing into Cells: The image is divided into small 8x8 cells. A histogram is created for each cell using the gradient information. This histogram shows the direction and magnitude of the gradients in that cell.

Creating the HOG Vector: The histograms from each cell are combined into larger regions (e.g., 16x16 pixels). A vector representing the features of this region is created. These feature vectors can be used in object recognition and machine learning algorithms.

The HOG descriptor's emphasis on gradient-based features and its robustness to changes in illumination make it a widely used technique in computer vision applications. Figure 3 demonstrates the HOG.



Fig. 3. Example of feature extraction step by step of HOG descriptor.

D. Multi task cascade convolutional neural network

MTCNN is a framework that combines face detection and face alignment into a unified approach using cascaded convolutional neural networks (CNNs) and multi-task learning. Traditional MTCNN [20] uses standard convolution for its operations. The framework consists of three key networks. The first, the Proposal Network (P-Net), generates candidate windows and their bounding box regression vectors, applying non-maximum suppression (NMS) to eliminate overlapping boxes. The second, the Refine Network (R-Net), filters out false positives from P-Net's

output and adjusts the bounding boxes through regression. Finally, the Output Network (O-Net) refines the final candidate windows and predicts the positions of five facial landmarks using a deeper network architecture. This integrated approach effectively handles both localization and alignment tasks, making MTCNN a popular choice for face detection applications. The illustration of key networks are demonstrated in Figure 4.

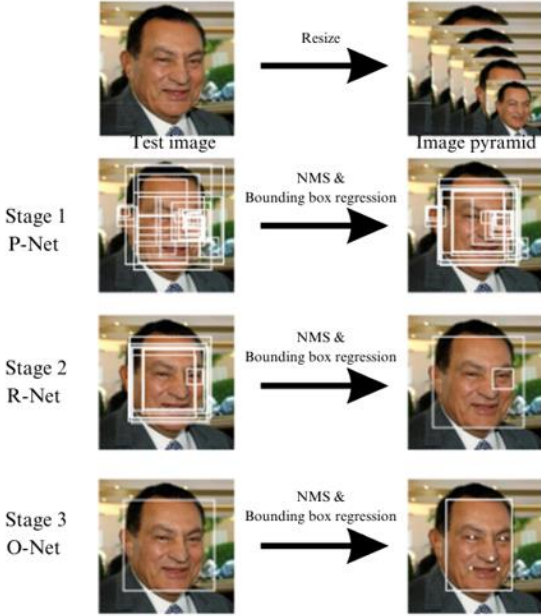


Fig. 4. Workflow of the MTCNN framework, demonstrating the three-stage process (P-Net, R-Net, O-Net) with image pyramid generation, non-maximum suppression (NMS), and bounding box regression for face detection and alignment [20].

III. EXPERIMENTAL RESULTS

The experimental was implemented using Python 3.11 on a Windows 11 Pro operating system, with Visual Studio Code as the development environment. TensorFlow and OpenCV libraries were utilized for implementation of each model. All the experiments were conducted on a high-performance gaming laptop with an 11th Gen Intel® Core™ i7-11800H CPU (2.30 GHz), 16 GB of RAM, and an NVIDIA GeForce RTX 3050 GPU with 6 GB memory. Note that, all analysis were performed on CPU.

The performance of each model is evaluated based on two criteria: accuracy and processing time. Processing time refers to the total detection time required by each method to process all images in the dataset and the average detection time per sample. Accuracy is calculated using the following metric:

$$\text{Acc} = \frac{(\text{Nr. of Detected Images})}{(\text{Nr. of Detected Images} + \text{Nr. of Undetected Images})} \times 100$$

Table I compares the performance of Haar Cascade, HOG, and MTCNN methods in terms of detected images, undetected images, and test accuracy. Among the methods, MTCNN demonstrates the highest test accuracy at 99.99%, with only one undetected image, showcasing its superior reliability for image detection tasks. Haar Cascade achieves a test accuracy of 99.57% but with 56 undetected images, indicating a

remarkable gap in precision compared to MTCNN. HOG, on the other hand, records a slightly lower test accuracy of 99.27% and 96 undetected images, reflecting its limitations in detection performance. This analysis highlights MTCNN as the most effective method for accurate and comprehensive image detection in this context.

TABLE I. Performance Analysis of Haar Cascade, HOG and MTCNN Methods in Terms of Accuracy.

Methods	Detected Images	Undetected Images	Test Accuracy (%)
Haar Cascade	13.177	56	99,57
HOG	13.137	96	99,27
MTCNN	13.232	1	99,99

Table II compares the processing time performance of Haar Cascade, HOG, and MTCNN methods in terms of total detection time and detection time per sample for a dataset of 13,233 images. Haar Cascade demonstrates the fastest processing time, with a total detection time of 178 seconds and an average detection time per sample of 13.45 microseconds. In contrast, HOG has the slowest performance, with a total detection time of 875 seconds and 66.12 microseconds per sample. MTCNN, while achieving higher accuracy (as shown in Table I), balances processing time with a total of 789 seconds and an average of 59.62 milliseconds per sample. These results indicate that while Haar Cascade is the most time-efficient method, MTCNN offers a trade-off between accuracy and processing speed, making it a viable option for applications requiring both performance metrics.

TABLE II. Performance Analysis of Haar Cascade, HOG and MTCNN Methods in Terms of Processing Time.

Methods	Total Images	Total Detection Time (sec)	Detection Time per sample (ms)
Haar Cascade	13.233	178	13.45
HOG	13.233	875	66.12
MTCNN	13.233	789	59.62

Figure 5 show cases the faces detected by HOG, Haar Cascade, and MTCNN methods, highlighting the detection success of each approach. Each method's detected faces are enclosed in distinct colored frames, visually representing their performance in identifying faces within the dataset. This figure emphasizes the differences in accuracy and detection reliability across the methods, with MTCNN appearing to achieve more consistent and precise detections.

Figure 6 illustrates examples of images that were not detected by HOG, Haar Cascade, and MTCNN methods. These undetected images highlight the limitations of each method, particularly Haar Cascade and HOG, in detecting certain faces, possibly due to variations in pose, lighting, or occlusions.

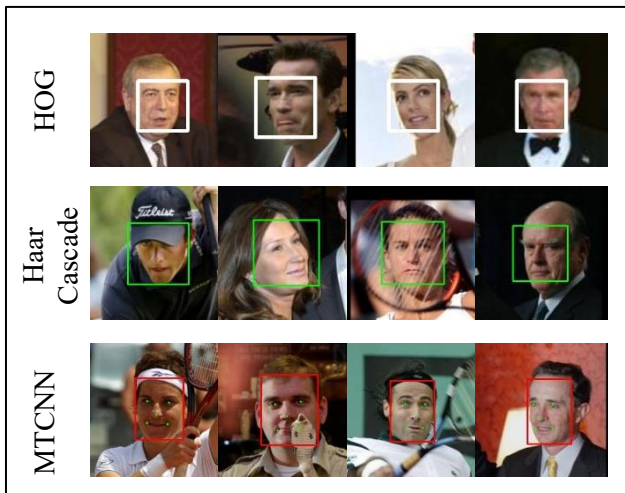


Fig. 5. Images detected during face detection with each method.

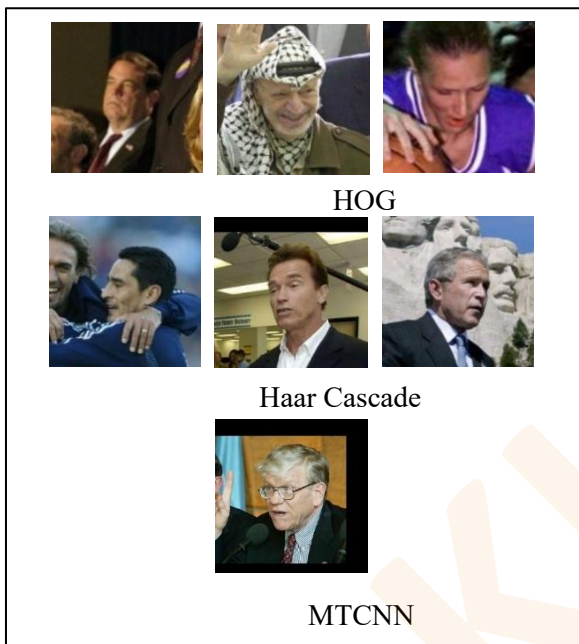


Fig. 6. Examples of undetected images of each methods.

IV. CONCLUSION

This study compares the Haar Cascade, HOG, and MTCNN methods for face detection using the LFW dataset. MTCNN emerges as the most accurate method, failing to detect only one image, while Haar Cascade and HOG demonstrate significantly higher failure rates. Although Haar Cascade is the fastest method in terms of processing time, it sacrifices accuracy, making it less reliable for complex detection tasks. HOG, while slightly more accurate than Haar Cascade, exhibits the longest processing time and the highest number of undetected faces. The balance between high accuracy and reasonable processing time makes MTCNN the most effective approach for face detection in real-world applications, especially where precision is critical. These findings underscore the importance of selecting methods based on application-specific requirements, such as accuracy and speed.

REFERENCES

[1] Z. Dong, J. Wei, X. Chen, and P. Zheng, "Face Detection in Security Monitoring Based on Artificial Intelligence Video

- Retrieval Technology," *IEEE Access*, vol. 8, pp. 63421–63433, 2020, doi: 10.1109/ACCESS.2020.2982779.
- [2] A. Khuran, B. P. Lohani, V. Bibhu, and P. K. Kushwaha, "An AI Integrated Face Detection System for Biometric Attendance Management," in *Proceedings of 2021 2nd International Conference on Intelligent Engineering and Management, ICIEM 2021*, Institute of Electrical and Electronics Engineers Inc., Apr. 2021, pp. 29–33. doi: 10.1109/ICIEM51511.2021.9445295.
- [3] Y. Shi, Z. Zhang, K. Huang, W. Ma, and S. Tu, "Human-computer interaction based on face feature localization," *J Vis Commun Image Represent*, vol. 70, Jul. 2020, doi: 10.1016/j.jvcir.2019.102740.
- [4] Y. Sun, X. Wang, and X. Tang, "Deep Learning Face Representation from Predicting 10,000 Classes."
- [5] F. Z. Canal *et al.*, "A survey on facial emotion recognition techniques: A state-of-the-art literature review," *Inf Sci (N Y)*, vol. 582, pp. 593–617, Jan. 2022, doi: 10.1016/j.ins.2021.10.005.
- [6] V. Raman, K. Elkarazle, and P. Then, "Gender-specific Facial Age Group Classification Using Deep Learning," *Intelligent Automation and Soft Computing*, vol. 34, no. 1, pp. 105–118, 2022, doi: 10.32604/iasc.2022.025608.
- [7] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2001, doi: 10.1109/CVPR.2001.990517.
- [8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 1, pp. 886–893, 2005, doi: 10.1109/CVPR.2005.177.
- [9] V. Vapnik, *The nature of statistical learning theory*. 2013.
- [10] M. Bertozzi, A. Broggi, M. Del Rose, M. Felisa, A. Rakotomamonjy, and F. Suard, "A Pedestrian Detector Using Histograms of Oriented Gradients and a Support Vector Machine Classifier." [Online]. Available: <http://www.tacom.army.mil>
- [11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
- [12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection." [Online]. Available: <https://goo.gl/bEs6Cj>.
- [13] W. Liu *et al.*, "SSD: Single shot multibox detector," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9905 LNCS, pp. 21–37, 2016, doi: 10.1007/978-3-319-46448-0_2/FIGURES/5.
- [14] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Process Lett*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.
- [15] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," 2008, Accessed: Dec. 06, 2024. [Online]. Available: <https://inria.hal.science/inria-00321923>
- [16] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," 2004.
- [17] A. B. Shetty, Bhoomika, Deeksha, J. Rebeiro, and Ramyashree, "Facial recognition using Haar cascade and LBP classifiers," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 330–335, Nov. 2021, doi: 10.1016/j.gltp.2021.08.044.
- [18] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features."
- [19] S. Kaur and D. Sharma, "Comparative Study of Face Detection Using Cascaded Haar, Hog and MTCNN Algorithms," in *2023 3rd International Conference on Advancement in Electronics and Communication Engineering, AECE 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 536–541. doi: 10.1109/AECE59614.2023.10428242.
- [20] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Process Lett*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

Design of a Circularly Polarized Antenna Operating at 137 MHz for NOAA Weather Satellite Reception

Suad Başbuğ
Department of Electrical Electronics Engineering
Nevşehir Hacı Bektaş Veli University
Nevşehir, Türkiye
suad@nevsehir.edu.tr

İrem Karataş Çetin
Department of Electrical Electronics Engineering
Nevşehir Hacı Bektaş Veli University
Nevşehir, Türkiye
ikaratas1997@gmail.com

Abstract— This study presents the design and simulation of a circularly polarized antenna operating at 137 MHz for receiving image data from NOAA satellites. The antenna design consists of two V-dipole antennas arranged orthogonally with a 90° phase difference to achieve circular polarization. Rod reflectors are added into design to enhance directivity. The final antenna design exhibits a realized gain of 4.82 dBi and an axial ratio of 0.51 dB in the maximum radiation direction. This antenna system offers an efficient solution for NOAA satellite signal reception, combining circular polarization with improved gain performance.

Keywords—NOAA satellites, circular polarization, V-dipole antenna, rod reflector

I. INTRODUCTION

National Oceanic and Atmospheric Administration (NOAA) satellites are in the category of meteorological and environmental observation satellites. NOAA system has five geostationary, two polar-orbiting and one deep space satellites. The primary task of NOAA satellites is to monitor weather, climate, and environmental conditions. These satellites are also used to mitigate the devastating effects of natural disasters such as hurricane, flood, wildfire, tornado, volcano and tsunami.

With the help of software-defined radio (SDR) technology, the works of directly capturing data from the earth observation satellites have found widespread use. In some works, quadrifilar helix (QFH) antennas are preferred because of their high gain and circular polarization characteristics [1]. Though the design and construction of QFH antennas are relatively harder than the classical dipole antennas, they are better alternatives compared to large parabolic antennas [2]. For the sake of achieving circular polarization, the design and fabrication of double cross antennas can be easier, but they still need baluns to work properly [3]. Using a balun for a single V-dipole antenna is not essential, but they have limited directivity capabilities [4]. Because of their circular polarization characteristics, turnstile antenna elements are also used to build a Yagi-Uda antenna for 137 MHz reception [5]. The double cross dipole antenna has good return loss values, but it has an omnidirectional radiation pattern [6]. The Moxon antenna is also analyzed in the literature and the study is concluded that it has relatively narrowband characteristics while its gain is sufficiently high [7].

In this study, a circularly polarized antenna operating at a frequency of 137 MHz is designed. For this antenna design, two V-dipole antennas are placed with a 90° phase difference to achieve circular polarization. Both V-dipole antennas in the same plane have the same basic geometry. The circular polarization is finally achieved by rotating one of the antennas with a 90° angle. In order to increase the directivity of the

antenna, a rod reflector is added for each V-dipole antenna element on the same plane with corresponding element. At every stage of the design, the S11 return loss and radiation pattern are monitored. The final design is a circularly polarized directional antenna working at 137 MHz. All simulations in this study are performed with the RF module of COMSOL Multiphysics software.

II. ANTENNA DESIGN AND SIMULATION RESULTS

The antenna design proposed in this study starts with a single basic V-dipole antenna model as shown in Fig. 1. The angle between the arms of the V-dipole antenna is 120°. Each arm of the antenna has a length of 540 mm. The cylindrical rods forming the arms have a radius of 1.382 mm. The overall structure of the antenna is positioned on the x-z plane while the feeding point of the arms located at the origin.

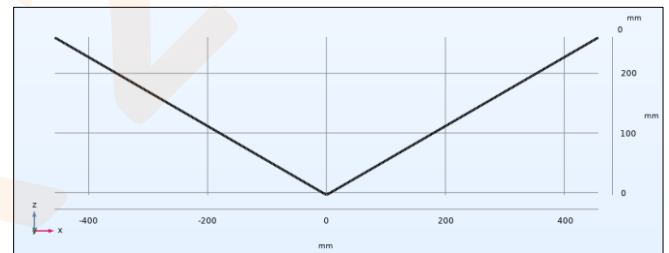


Fig. 1. Basic V-dipole antenna geometry

The simulated S11 plot for the antenna in this phase is given in Fig. 2. As can be seen from the S11 plot, a value of -18 dB at 137 MHz is achieved.

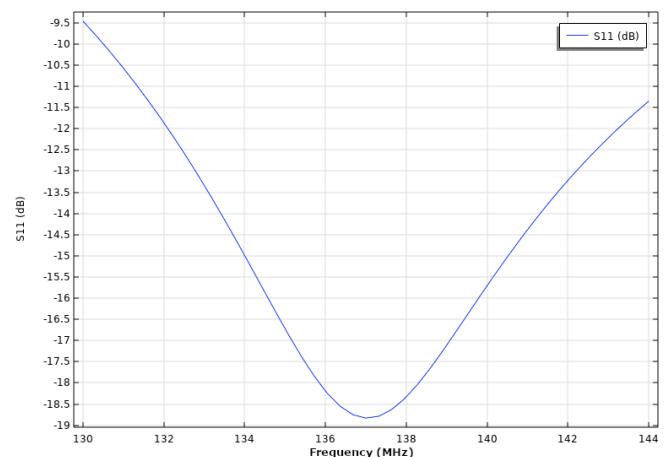


Fig. 2. S11 plot of single V-dipole antenna

Three-dimensional radiation pattern of the current design is presented in Fig. 3. The radiation pattern in the figure is very

similar to that of the classical dipol. The maximum realized gain of the single V-dipole antenna is 1.41 dBi.

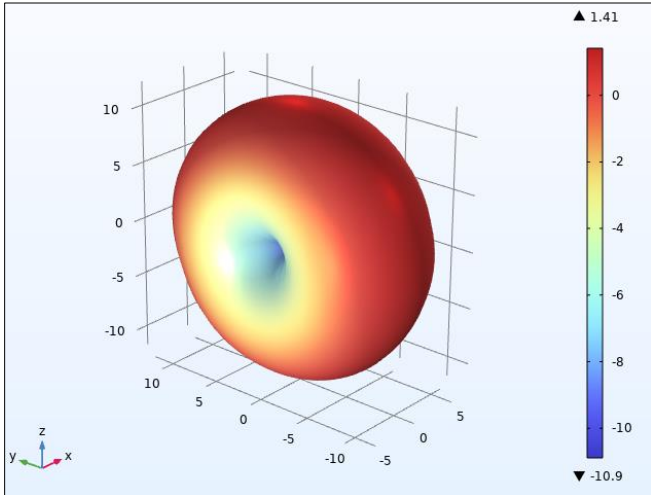


Fig. 3. Three-dimensional radiation pattern of single V-dipole antenna

In order to achieve circular polarization, another V-dipole having the same geometrical properties with the first one is placed in the overall antenna structure. The second V-dipole is placed 547 mm away from the first one as shown in Fig. 4. The second requirement for achieving circular polarization has been achieved by rotating the upper V-dipole by 90° around the z-axis. We assume that two antennas are fixed along the z-axis using an insulating rod material.

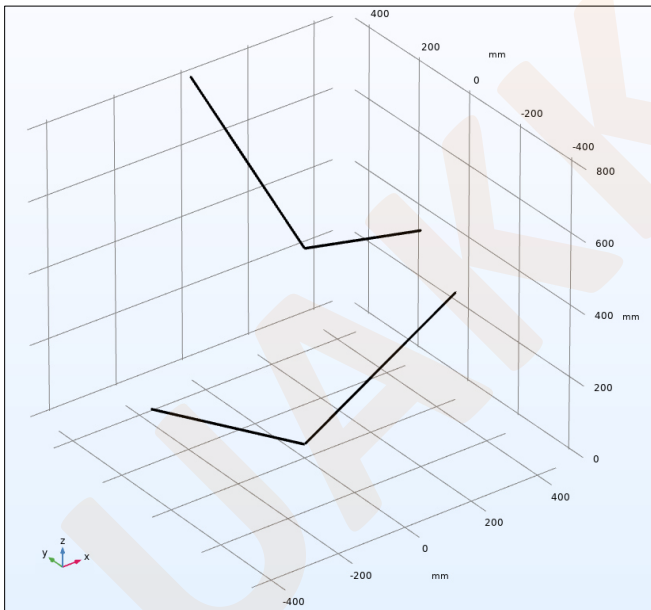


Fig. 4. Cross-positioned V-dipole antenna pair

The antennas are fed through separate ports in the simulation. The S11 parameters obtained for the lower and upper V-dipole antennas are given in Fig. 5. It can be seen from the figure that both S11 results have resonances very close to the target frequency of 137 MHz. Additionally, both plots successfully drop below -17 dB.

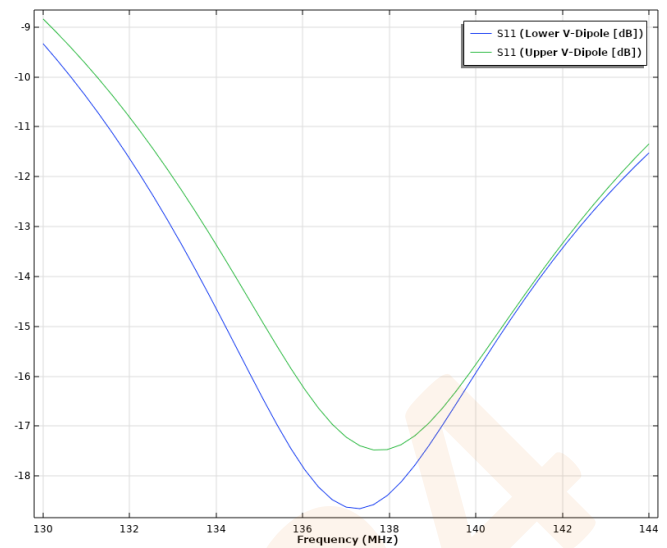


Fig. 5. S11 plots for the cross-positioned V-dipole antennas

In Fig. 6, the three-dimensional radiation pattern of the dual V-dipole antenna system is presented. At this stage of the design, the maximum realized gain value is 1.24 dBi. These results indicate that the antenna system has partially shifted from an omnidirectional structure to a somewhat directional antenna. Accordingly, it can be observed that the radiation is concentrated in both the positive and negative z-directions.

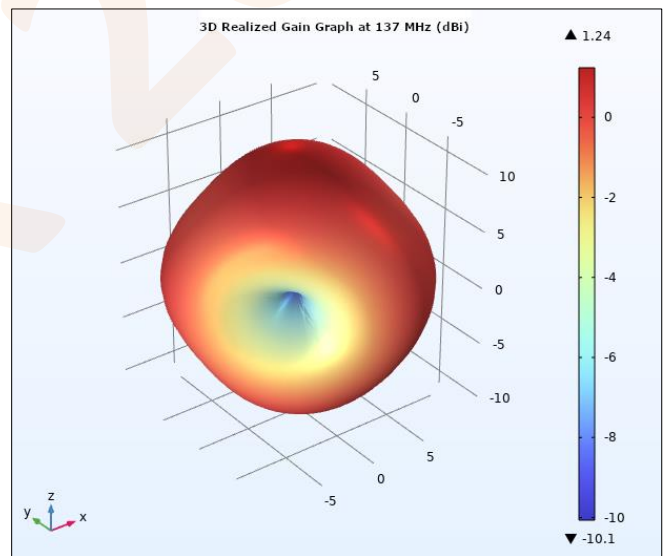


Fig. 6. Three-dimensional radiation pattern of the V-dipole antenna pair

The axial ratio of the V-dipole antenna pair at 137 MHz is presented in Fig. 7. As can be clearly seen from the figure, the axial ratio value towards the z-axis is significantly lower than 3 dB. An axial ratio value of 0.15 dB has been successfully achieved in the positive z-axis direction.

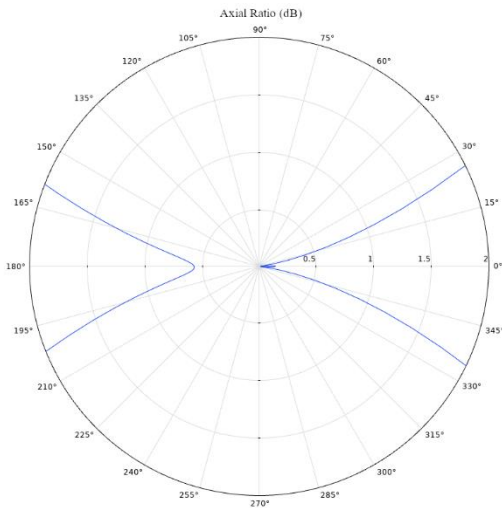


Fig. 7. Axial ratio plot of the V-dipole antenna pair at a 137 MHz center frequency

After confirming that the results of the circularly polarized dual V-dipole antenna design meet the requirements, we begin to improve the directivity of the antenna system. At this stage, a rod reflector is added to each V-dipole antenna element, as shown in Fig. 8. The distance between the V-dipole antennas and their respective reflectors is set to 450 mm. Each reflector is a cylindrical rod with a radius of 4 mm and a length of 1100 mm.

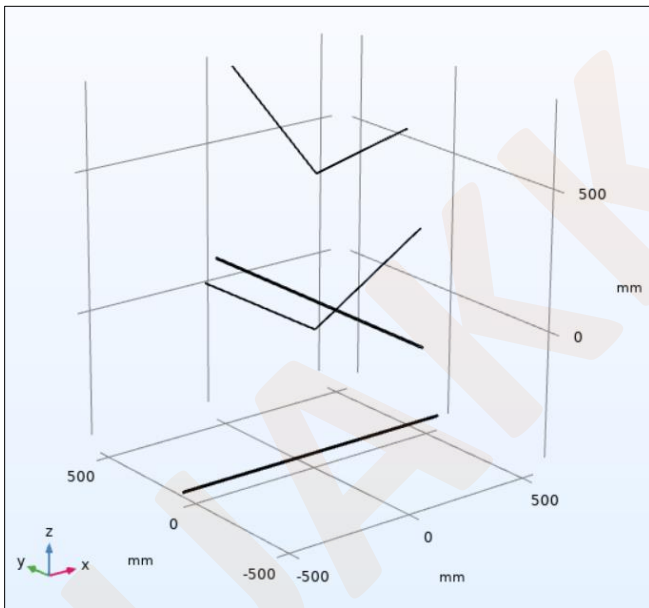


Fig. 8. V-dipole antenna pair with rod reflectors

Rod reflectors added to enhance directivity cause the resonance point in the S11 graphs to shift left along the frequency axis and rise along the vertical axis. In order to compensate this change, the arm lengths of V-dipoles are reconsidered. As a result of the parametric sweep studies, the arm lengths are reduced to 515 mm. The new S11 graphs obtained from the simulations are presented in Fig. 9. The S11 value in both plots at 137 MHz is below -13 dB.

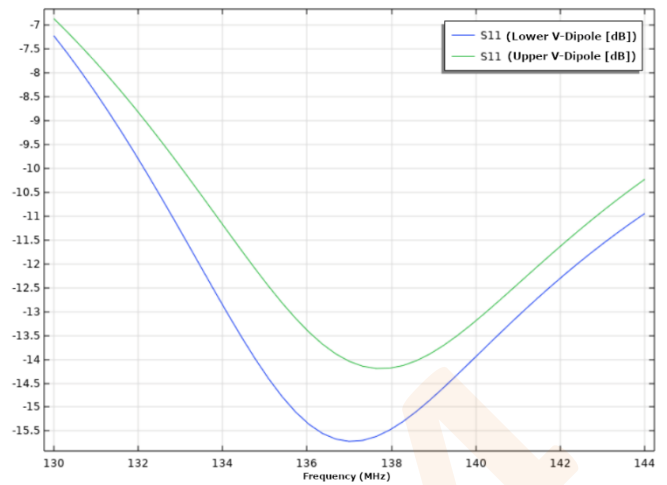


Fig. 9. S11 plots of the V-dipole antenna pair with rod reflectors

The three-dimensional radiation pattern of V-dipole antenna pair with rod reflectors is presented in Fig. 10. The positive effect of the reflectors on gain is clearly seen from the figure. This diagram is generated based on the realized gain values obtained from the simulations of the designed antenna system. The maximum realized gain has increased to a level of 4.82 dBi.

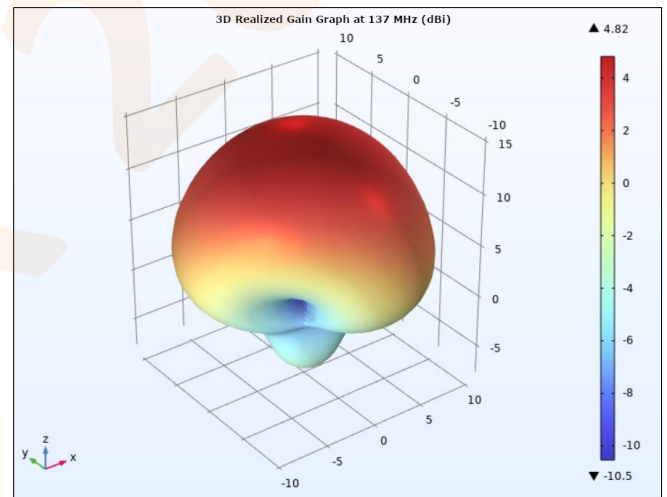


Fig. 10. Three-dimensional radiation pattern of the V-dipole antenna pair with rod reflectors

Finally, the axial ratio graph of the final design is presented in Fig. 11. The axial ratio in the positive z-axis direction is 0.51 dB, which is significantly lower than the upper limit of 3 dB. Additionally, it is observed that the axial ratio successfully remains below 3 dB over a wide angle of 60° in the x-z plane.

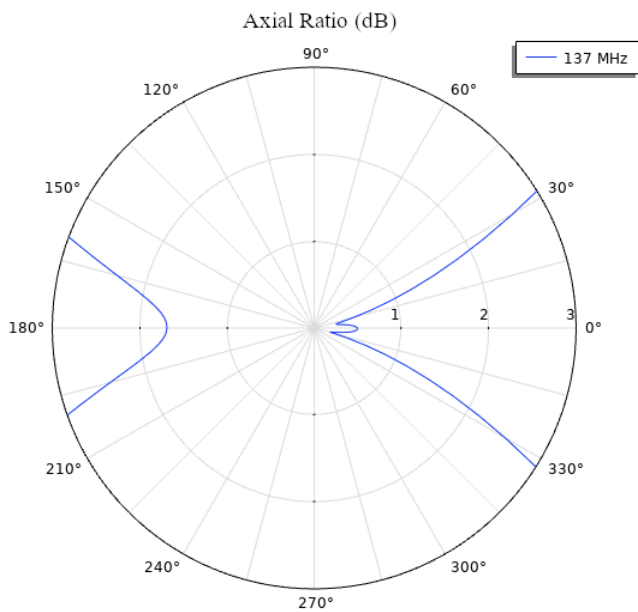


Fig. 11. Axial ratio plot of the V-dipole antenna pair with rod reflectors at a 137 MHz center frequency

III. CONCLUSION

In this study, a circularly polarized antenna operating at 137 MHz is designed and simulated to be used for receiving image data from NOAA satellites. The proposed antenna design utilizes two V-dipole antennas placed orthogonally with a 90° phase difference to achieve circular polarization. Rod reflectors are added to enhance the directivity of antenna design. Adjustments to the lengths of the V-dipole arms ensure that the antenna resonance is maintained at the target frequency after the reflectors are added. The final design has a realized gain of 4.82 dBi and an axial ratio of 0.51 dB toward

the positive z-axis. The proposed antenna system provides an efficient solution for receiving NOAA satellite signals, combining circular polarization and improved gain.

REFERENCES

- [1] S. Mahmood, M. T. Mushtaq, and G. Jaffer, "Cost efficient design approach for receiving the NOAA weather satellites data," in *2016 IEEE Aerospace Conference*, Mar. 2016, pp. 1–6.
- [2] R. Wiryadinata, M. Khoirussolih, N. Rohanah, I. Muttakin, and T. Firmansyah, "Image Data Acquisition for NOAA 18 and NOAA 19 Weather Satellites Using QFH Antenna and RTL-SDR," *MATEC Web Conf.*, vol. 218, p. 02002, 2018.
- [3] F. P. A. Escobedo, H. R. Álvarez, H. Salazar, C. G. R. Percing, and R. L. J. M. de Oca, "Low cost optimization method of a double cross antenna satellite reception system for the processing and improvement of meteorological satellite signals and images NOAA 15-18-19," in *2019 IEEE 1st Sustainable Cities Latin America Conference (SCLA)*, Aug. 2019, pp. 1–6.
- [4] M. Fathurahman, Zulhelman, A. Maulana, and M. Widyawati, "Design and Development of Dipole Antenna for NOAA Satellite Image Acquisition System and Processing," *J. Phys. Conf. Ser.*, vol. 1364, no. 1, p. 012025, Dec. 2019.
- [5] G. Ferré, A. Ghiotto, and E. Kerhervé, "Design and Deployment of a Ground Station to Demodulate NOAA Meteorologic Images," in *2019 29th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEIE)*, Sep. 2019, pp. 1–10.
- [6] Y. Rafsyam, I. Z. E. E. Khairas, Jonifan, and W. A. Karimah, "Design of Double Cross Dipole Antenna as NOAA Satellite Signal Receiver for Monitor Cloud Conditions Application," *J. Phys. Conf. Ser.*, vol. 1364, no. 1, p. 012059, Dec. 2019.
- [7] H. P. Penagos, E. M. Mahecha, A. P. Avendaño, and R. C. Riaño, "Design, Construction, Evaluation and Comparison of Antennas for Reception of NOAA Signals," *Cienc. Poder Aéreo*, vol. 19, no. 2, Art. no. 2, Jul. 2024.

Comparison of Large Language Models (LLM)

Yuksel Celik

*Information Security and Digital Forensics
University at Albany, State University of New York Albany
NY, USA
ycelik@albany.edu*

Sakshi Singh

*College of Emergency Preparedness, Homeland Security and
Cybersecurity
University at Albany, State University of New York, Albany
NY, USA
ssingh29@albany.edu*

Lakshika Lakshika

*College of Emergency Preparedness, Homeland Security and
Cybersecurity
University at Albany, State University of New York, Albany
NY, USA
lvaishnav@albany.edu*

Sanjay Goel

*Information Security and Digital Forensics
University at Albany, State University of New York Albany
NY, USA
goel@albany.edu*

Abstract— The rapid advancement of artificial intelligence (AI) has led to the emergence of powerful Large Language Models (LLMs), revolutionizing the field by enabling advanced language understanding and generation capabilities. This study conducts a comprehensive comparative analysis of prominent LLMs, including OpenAI's ChatGPT, Google's Gemini, Meta's LLaMA, and Hugging Face's BLOOM. The analysis evaluates these models based on their architecture, training methodologies, performance, application areas, and unique features. While ChatGPT excels in language generation, Gemini offers multimodal capabilities, LLaMA provides open-source flexibility for academic research, and BLOOM stands out for its customization options. The findings provide critical insights into selecting and optimizing LLMs for diverse applications while addressing their limitations. This work aims to serve as a valuable resource for researchers, developers, and industry professionals navigating the rapidly evolving landscape of LLM technology.

Keywords— Large Language Model, ChatGPT, LLaMA, Gemini, BLOOM

I. INTRODUCTION

The emergence of deep learning methods in the field of artificial intelligence (AI) has opened up opportunities for AI applications across numerous domains. However, the rapid development of Large Language Models (LLMs) and their exceptional capabilities has marked a transformative milestone in the evolution of AI.

LLMs, with their advanced features, have not only simplified many tasks in daily life but have also found extensive use in scientific research[1], healthcare[2][3][4][5], law[6], materials science[7], biology[8], education[9], software development[10], autonomous systems[11], and manufacturing[12]. OpenAI's ChatGPT[13] was the first LLM to be introduced, bringing widespread adoption due to its creative text generation and strong logical reasoning capabilities. Subsequently, Google developed Gemini[14], which is distinguished by its multimodal capabilities,

processing textual and visual data simultaneously, making it particularly effective in applications such as search engines and creative assistants. Meta followed with the development of the LLaMA[15] models, which adopt an open-source approach to deliver innovative solutions, especially for the academic and research communities.

In addition to these three major models, IBM's WatsonX platform caters to corporate needs by providing tailored AI solutions, while Hugging Face's community-supported open-source models, such as BLOOM[16] and Falcon[17], have become popular for research and development projects.

This study aims to provide an in-depth analysis of the technical capacities, innovative features, and application areas of prominent large language models. It will evaluate which models are most suitable for specific use cases by analyzing their strengths and weaknesses. This comprehensive comparison seeks to provide a better understanding of the current state of AI technologies and their potential future trajectories. The analysis serves as a guide for researchers and practitioners, facilitating the process of selecting the most appropriate AI solution.

II. LARGE LANGUAGE MODELS (LLMs)

A. General Framework of LLMs

Large Language Models (LLMs) are deep learning-based models trained on extensive datasets to acquire human-like abilities in understanding, generating, and manipulating language. Their fundamental working principles are outlined in the following stages:

1) Transformer Architecture

Most LLMs are built upon the Transformer architecture, as illustrated in Figure 1 [18]. The primary strength of this architecture lies in its utilization of the attention mechanism.

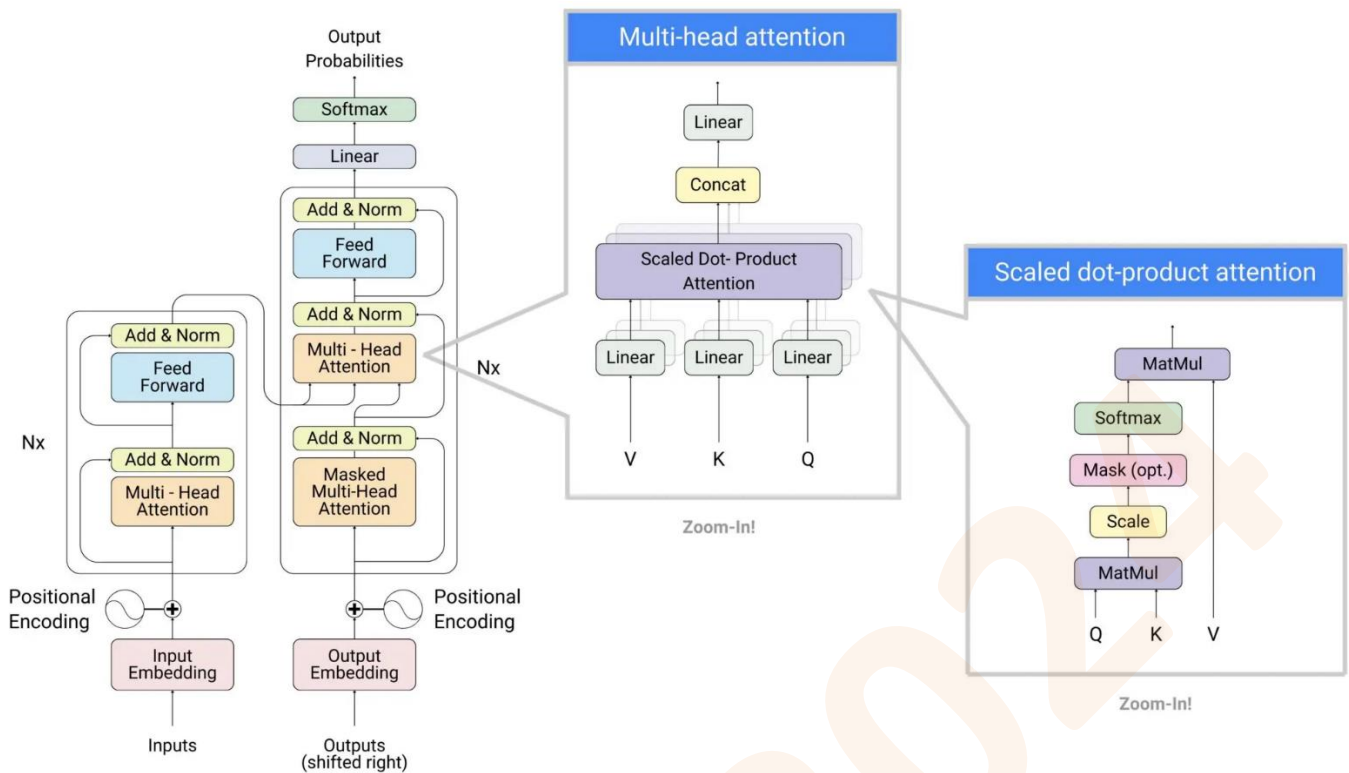


Fig. 1 The Transformer Architecture and Multi-Head attention mechanism

Attention Mechanism: This neural network architecture enables a deep learning model to focus on specific and relevant aspects of the input data. It allows machines to better understand the input and generate appropriate outputs. The workflow includes:

Self-Attention: Enables each word (or token) in a sentence to learn its relationship with other words, allowing for meaningful contextual relationships to be established.

Multi-Head Attention: Facilitates parallel processing of attention mechanisms across different "heads," enabling the model to learn various contexts simultaneously.

Positional Encoding: Since Transformers work with sequential data, positional encoding adds position information to tokens to help the model learn the order of words.

2) LLM Training Process

The initial steps in training LLMs involve data processing and tokenization:

Data Collection: Models are typically trained on a wide variety of textual data, including web pages, books, datasets, articles, and code.

Tokenization: Text is divided into words or subword units (tokens), using algorithms like Byte Pair Encoding (BPE) [19] or SentencePiece [20].

Language Modeling Tasks:

Learning Probability Distributions: The model learns the probability distribution of a token within its context.

Causal Language Modeling (CLM): Predicts future tokens (e.g., GPT models).

Masked Language Modeling (MLM): Predicts randomly masked tokens (e.g., BERT).

Loss Function: Cross-entropy loss is typically used to minimize the difference between the model outputs and true values.

3) Model Parameter Size

LLMs can contain hundreds of billions to trillions of parameters. The larger the number of parameters, the better the performance. These parameters include:

Weight Matrix: Learns the relationships between words and contexts.

Number of Layers: Deeper models can learn more complex relationships.

Hidden Size: Determines the information processing capacity within each layer.

4) Prediction in LLMs

Once training is complete, the model can make predictions on new data:

Input: A sequence of tokens is fed into the model and converted into numerical vectors in the embedding layer.

Contextual Processing: Inputs are processed through Transformer layers to learn contextual relationships.

Output Generation: At each step, the model generates a token probability distribution and selects the most likely token.

5) Optimization Techniques

Efficient training of large models involves the following techniques:

Distributed Training: Distributes data and the model across multiple processors (GPU/TPU) for parallel processing.

Mixed Precision: Combines 16-bit and 32-bit operations to reduce memory and computation costs.

Fine-Tuning: Trains the model on a broad dataset for general knowledge and then fine-tunes it on specific tasks with specialized data.

III. MOST POPULAR LLMs AND COMPARISONS

A. ChatGPT

OpenAI's ChatGPT series (GPT-3, GPT-3.5, GPT-4) stands out for its broad language understanding and creative content generation capabilities. These models are widely utilized in various areas, including chatbots, content creation, code writing, and summarization. With a robust API ecosystem, ChatGPT can be seamlessly integrated into different platforms, addressing the needs of users across a wide range of languages [21].

B. Google DeepMind (Gemini/AlphaCode)

Google's Gemini and AlphaCode models offer advanced capabilities in language understanding, code generation, and multimodal (text and visual) processing. These models are particularly useful for search engine enhancements, creative assistants, and scientific research. Their strong integration with Google's product ecosystem provides users with both functional and creative solutions [22].

C. Meta (LLaMA Series)

Meta's LLaMA models (LLaMA 1 and LLaMA 2) are open-source AI models designed for research purposes. These models are well-suited for academic projects and community-driven development efforts. Their open access to the research community allows for continuous improvement by contributors [23].

D. Hugging Face (BLOOM)

BLOOM, developed by Hugging Face, is notable for its open-source and customizable nature. It is a popular choice for research and development projects, supported by a large developer community. The flexibility of customization makes it applicable to a wide variety of projects [24].

While the general working principles of LLMs are similar, they differ structurally, resulting in unique advantages and disadvantages depending on the application. These differences are highlighted in the following tables: Architectural Structures: Table 1, Model Training Processes: Table 2, Performance and Competencies: Table 3, Usage Scenarios: Table 4, Chronological Development and Improvements: Table 5, Languages Used in Training Data: Table 6

TABLE 1. COMPARISON ACCORDING TO ARCHITECTURAL STRUCTURES

Model	Basic Architecture	Model Architecture	Attention Mechanism	Input Data
ChatGPT-4 (OpenAI)	Transformer base	175+ billion parameters	CLM	Text
Gemini (Google)	Transformer + Multimodal abilities	1 trillion parameters	CLM+ Multimodal	Text, Visual, Table
LLaMA (Meta)	Transformer base	7B, 13B, 65B (different size)	CLM	Text

BLOOM (Hugging Face)	Transformer base	176 billion parameters	CLM	Text
----------------------	------------------	------------------------	-----	------

In Table 1, a comparison based on architectural structures reveals that all models utilize the Transformer architecture at their core. Their parameter counts exceed billions, and they employ Causal Language Modeling (CLM) as the attention mechanism. Additionally, text is observed to be the primary input data for all models.

Gemini stands out from the other models with its multimodal capabilities in both architecture and attention mechanisms. Unlike the others, Gemini can process not only textual data but also visual data and tables as input, making it uniquely suited for diverse applications.

TABLE 2. MODEL TRAINING PROCESSES

Model	Training Data	Data Processing	Training Method	Objective Function
ChatGPT-4 (OpenAI)	A large text dataset, books, web pages	Owned and optimized data processing	Pre-training + Fine-tuning (with RLHF)	Cross Entropy Loss + Human Feedback (RLHF)
Gemini (Google)	Multimodal data (text + images)	Specialized processing for multimodality	Pre-training + Fine-tuning	Cross Entropy Loss + Visual Loss
LLaMA (Meta)	Academic articles, internet data	Academic-oriented data filtering	Pre-training	Cross Entropy Loss
BLOOM (Hugging Face)	Open datasets	Open data and community-based processing	Pre-training	Cross Entropy Loss

An examination of Table 2 reveals that, in terms of training methodology, ChatGPT-4 and Gemini distinguish themselves by incorporating fine-tuning in addition to their pre-training processes. This additional fine-tuning step allows these models to achieve enhanced performance and adaptability for specific tasks compared to models that rely solely on pre-training.

TABLE 3. PERFORMANCE AND COMPETENCIES

Model	Natural Language Generation	Code Generation	Multimodality	Speed and Efficiency
ChatGPT-4 (OpenAI)	Highly successful in producing human-like text	Advanced level in coding tasks	Text-based only	High accuracy, slower
Gemini (Google)	Rich content production with multimodality capabilities	Code generation is strong but not as strong as GPT	Ability to process text, images, tables	Very fast, powerful in context combined with visuals

LLaMA (Meta)	Strong understanding of text production and context	Limited in code generation	Text-based only	Lighter and faster
BLOOM (Hugging Face)	Text production and open source contribution	Intermediate code generation	Text-based only	Good performance on large data sets

As per Table 3, a comparison of natural language processing performance indicates that ChatGPT-4 stands out with superior performance, showcasing its ability to generate human-like text more effectively than its counterparts.

TABLE 4. AREAS OF USAGE

Model	Chat Apps	Search Engine Ability	Scientific Research	Open Source
ChatGPT-4 (OpenAI)	Yes	Limited	Medium level	No
Gemini (Google)	Yes	Google search integration	Powerful	No
LLaMA (Meta)	Yes	Limited	Academically oriented	Yes
BLOOM (Hugging Face)	Yes	No	Research-oriented	Yes

According to Table 4, all models support chat applications as a common usage scenario. However, in terms of search engine capabilities, Gemini stands out due to its integration with Google's search engine, providing a significant advantage in this area. Conversely, BLOOM lacks search engine functionality, limiting its applications in this domain.

TABLE 5. LLM MODELS CHRONOLOGICAL BREAKDOWN OF THE DEVELOPMENT AND IMPROVEMENTS

Model	2021	2022	2023	2024
ChatGPT-4 (OpenAI)	GPT-1,2,3	ChatGPT 3.5	GPT-4	GPT-4-o
Gemini (Google)	LaMDA	Pathways	Gemini 1	Gemini 1.5
LLaMA (Meta)	n/a	n/a	LLaMA 1	LLaMA 2
BLOOM (Hugging Face)	n/a	BLOOM	BLOOM	BLOOM

TABLE 6. LLM MODELS LANGUAGES INCLUDED IN THE TRAINING DATA

Model	Approx Languages	Description
ChatGPT-4 (OpenAI)	Almost all languages	Best for English, supports major global languages
Gemini (Google)	100+	Extensive multilingual support via Google datasets

LLaMA (Meta)	20+	Widely spoken languages, with less focus on niche ones
BLOOM (Hugging Face)	46	Multilingual inclusivity, including low-resource languages

According to Table 6, a comparison of the languages included in the training data shows that ChatGPT-4 ranks first with the most extensive language support, followed by Gemini.

IV. DISCUSSION

Although LLMs utilize similar AI methodologies, their processed data sources and structural differences result in distinct advantages and disadvantages depending on the intended use case.

ChatGPT stands out as a leader due to its strong language generation capabilities, enabling superior human-like text production, broad usage scenarios, and user-friendly design. However, its lack of multimodal capabilities and high operational costs are considered its key weaknesses.

Gemini benefits from seamless integration with the Google ecosystem and its multimodal capabilities, enabling the combination of various data types to support innovative applications. Despite these strengths, its weaknesses include being a closed-source model and having limited accessibility.

LLaMA, with its open-source and research-oriented design, is a powerful tool for academic environments. However, its limited training data and weaker performance in practical applications are notable disadvantages.

BLOOM, supported by Hugging Face's community-driven open-source approach, offers extensive datasets and strong customization capabilities, making it highly flexible for different projects. Despite these strengths, its performance does not match that of GPT or Gemini, and it requires significant resources, which are key limitations.

V. CONCLUSION

Large Language Models (LLMs) have become pivotal in advancing artificial intelligence, offering diverse capabilities that cater to a broad spectrum of applications. This study highlights the distinct strengths and weaknesses of leading LLMs, such as ChatGPT's unparalleled natural language processing performance, Gemini's multimodal abilities, LLaMA's open-source research orientation, and BLOOM's community-driven customization potential. Despite their shared foundation in Transformer architecture, these models exhibit unique features that make them suitable for specific use cases. However, challenges such as resource demands, accessibility limitations, and scalability must be addressed for their broader adoption. By providing a detailed comparative analysis, this research equips stakeholders with the knowledge to select the most appropriate model for their needs while contributing to the ongoing innovation and responsible deployment of LLMs.

REFERENCES

- [1] S. Nerella *et al.*, “Transformers and large language models in healthcare: A review,” *Artif Intell Med*, vol. 154, p. 102900, Aug. 2024, doi: 10.1016/j.artmed.2024.102900.
- [2] L. Verlingue, C. Boyer, L. Olgiati, C. Brutti Mairesse, D. Morel, and J. Y. Blay, “Artificial intelligence in oncology: ensuring safe and effective integration of language models in clinical practice,” *The Lancet Regional Health - Europe*, vol. 46, p. 101064, Nov. 2024, doi: 10.1016/J.LANEPE.2024.101064.
- [3] A. A. Birkun and A. Gautam, “Large Language Model-based Chatbot as a Source of Advice on First Aid in Heart Attack,” *Curr Probl Cardiol*, vol. 49, no. 1, p. 102048, Jan. 2024, doi: 10.1016/J.CPCARDIOL.2023.102048.
- [4] H. Hwai, Y. J. Ho, C. H. Wang, and C. H. Huang, “Large language model application in emergency medicine and critical care,” *Journal of the Formosan Medical Association*, Aug. 2024, doi: 10.1016/J.JFMA.2024.08.032.
- [5] R. Bommasani *et al.*, “On the Opportunities and Risks of Foundation Models,” Aug. 2021, Accessed: Dec. 16, 2024. [Online]. Available: <http://arxiv.org/abs/2108.07258>
- [6] G. Lei, R. Docherty, and S. J. Cooper, “Materials science in the era of large language models: a perspective,” *Digital Discovery*, vol. 3, no. 7, pp. 1257–1272, Jul. 2024, doi: 10.1039/D4DD00074A.
- [7] M. Bhattacharya, S. Pal, S. Chatterjee, S. S. Lee, and C. Chakraborty, “Large language model to multimodal large language model: A journey to shape the biological macromolecules to biological sciences and medicine,” *Mol Ther Nucleic Acids*, vol. 35, no. 3, p. 102255, Sep. 2024, doi: 10.1016/J.OMTN.2024.102255.
- [8] M. Haman and M. Školník, “Using ChatGPT to conduct a literature review,” *Account Res*, Dec. 2023, doi: 10.1080/08989621.2023.2185514/ASSET//CMS/ASSET/02DB9456-4A9F-4BE0-A4D1-80AE8E1B5F2E/08989621.2023.2185514.FP.PNG.
- [9] T. Alqahtani *et al.*, “The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research,” *Research in Social and Administrative Pharmacy*, vol. 19, no. 8, pp. 1236–1242, Aug. 2023, doi: 10.1016/J.SAPHARM.2023.05.016.
- [10] R. A. Husein, H. Aburajouh, and C. Catal, “Large language models for code completion: A systematic literature review,” *Comput Stand Interfaces*, vol. 92, p. 103917, Mar. 2025, doi: 10.1016/J.CSI.2024.103917.
- [11] L. Wang *et al.*, “A survey on large language model based autonomous agents,” *Front Comput Sci*, vol. 18, no. 6, pp. 1–26, Dec. 2024, doi: 10.1007/S11704-024-40231-1/METRICS.
- [12] S. Colabianchi, F. Costantino, and N. Sabetta, “Assessment of a large language model based digital intelligent assistant in assembly manufacturing,” *Comput Ind*, vol. 162, p. 104129, Nov. 2024, doi: 10.1016/J.COMPIND.2024.104129.
- [13] B. Li, V. L. Lowell, C. Wang, and X. Li, “A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT’s use in language learning and teaching,” *Computers and Education: Artificial Intelligence*, vol. 7, p. 100266, Dec. 2024, doi: 10.1016/J.CAEAI.2024.100266.
- [14] M. Masalkhi, J. Ong, E. Waisberg, and A. G. Lee, “Google DeepMind’s gemini AI versus ChatGPT: a comparative analysis in ophthalmology,” *Eye 2024* 38:8, vol. 38, no. 8, pp. 1412–1417, Feb. 2024, doi: 10.1038/s41433-024-02958-w.
- [15] H. Touvron *et al.*, “LLaMA: Open and Efficient Foundation Language Models,” Feb. 2023, Accessed: Dec. 16, 2024. [Online]. Available: <https://arxiv.org/abs/2302.13971v1>
- [16] T. Le Scao *et al.*, “BLOOM: A 176B-Parameter Open-Access Multilingual Language Model,” Nov. 2023, Accessed: Dec. 16, 2024. [Online]. Available: <https://inria.hal.science/hal-03850124>
- [17] V. Agatha and I. Setyawan, “Web Chat-based Application with Large Language Model and Transformers from Hugging Face for Self-Learning on Storytelling Skills,” *2024 International Electronics Symposium: Shaping the Future: Society 5.0 and Beyond, IES 2024 - Proceeding*, pp. 614–618, 2024, doi: 10.1109/IES63037.2024.10665795.
- [18] A. Vaswani *et al.*, “Attention is All you Need,” *Adv Neural Inf Process Syst*, vol. 30, 2017.
- [19] T. Xu and P. Zhou, “Feature Extraction for Payload Classification: A Byte Pair Encoding Algorithm,” *2022 IEEE 8th International Conference on Computer and Communications, ICC3 2022*, pp. 2441–2445, 2022, doi: 10.1109/ICCC56324.2022.10065977.
- [20] S. Choo and W. Kim, “A study on the evaluation of tokenizer performance in natural language processing,” *Applied Artificial Intelligence*, vol. 37, no. 1, Dec. 2023, doi: 10.1080/08839514.2023.2175112.
- [21] S. S. Gill and R. Kaur, “ChatGPT: Vision and challenges,” *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 262–271, Jan. 2023, doi: 10.1016/J.IOTCPS.2023.05.004.
- [22] R. Islam and I. Ahmed, “Gemini-the most powerful LLM: Myth or Truth,” *2024 5th Information Communication Technologies Conference, ICTC 2024*, pp. 303–308, 2024, doi: 10.1109/ICTC61510.2024.10602253.
- [23] J. Yeom *et al.*, “Tc-llama 2: fine-tuning LLM for technology and commercialization applications,” *J Big Data*, vol. 11, no. 1, pp. 1–31, Dec. 2024, doi: 10.1186/S40537-024-00963-0/TABLES/10.
- [24] B. Workshop *et al.*, “BLOOM: A 176B-Parameter Open-Access Multilingual Language Model Major Contributors Prompt Engineering Architecture and Objective Engineering Evaluation and Interpretability Broader Impacts,” 2023.

An Optimized Traffic Lights Scheduling

1st Oğuzhan Çiftçi

Dept. of Electrical & Electronics Engineering

Marmara University

Istanbul, Türkiye

0000-0002-9763-3677

Abstract—In this paper, we study on the optimal scheduling problem of the traffic lights on a map by utilizing the number of vehicles that use predefined routes. This paper considers modeling an optimization problem to minimize the number of vehicles on the map. We present the analytical solution of the optimization problem and the numerical solution approach by using an Open-Source library CVX implemented in Matlab. Also, we discuss the applications of our approach.

Keywords—constrained optimization, traffic light scheduling

I. INTRODUCTION

Traffic congestion is one of the most important problems in urban transportation [1, 2]. Especially the cost of time and money that caused by congestion have significant effects on human lives. It was estimated that an average person loses over 100 hours in traffic jams every year. Also, the extra money spent on fuel and the value of time combined was more than \$1000 per person in 2017 [3–5]. Besides, it can be argued that traffic jam has psychological effects on humans as well. Traffic jams may cause anger, high levels of stress, and behavioral anomalies for passengers as well as drivers [6–8].

Identifying the main causes of traffic congestion in urban areas is the primary step to relief the adverse effects of it. The structure of the roads is one of the most important indicators in traffic management. For example, the structure of the roads had to be designed by the boundaries of the buildings in central areas of old European cities. On the other hand, new types of transportation devices cannot be directly integrated into these conventional roads. Another significant indicator is human error. We may not be able to completely eradicate the human errors in traffic, however, we can supervise it by utilizing smart management approaches [3].

When we consider all these effects, we think that smart traffic scheduling solutions should be determined to ease the traffic conditions on roads. However, we need to define the traffic jam problem mathematically in order to solve it. We utilize the constrained optimization approach to find an optimal solution to traffic scheduling problem. However, we need to define some indicators to measure traffic congestion. There are several metrics have been articulated in [9, 10] such as travel time, speed, density, queue and distance. Nevertheless, these metrics also determine some of the constraints on the roads.

In this paper, we mainly focus on designing a constrained optimization problem by considering two types of road struc-

tures that can be used a basis to form more complex structures. We also discuss the analytical and numerical methods to reach an optimal solution of our problem. However, implementing an optimal solution on roads is as significant as finding a solution itself. Thus, we present some suggestions on applying our approach on urban roads.

This paper is organized as follows. In the following section, we provide a mathematical preliminary to form the problem of interest as a constrained optimization problem. In the next section, we introduce traffic flow structures that can be utilized in roads. In Section IV, we design a constrained optimization problem in accordance with the map type. In Section V, we present both the analytical and numerical solutions of the suggested optimization problem. In the last section, we conclude and discuss the main findings provided in this paper.

II. MATHEMATICAL PRELIMINARY

Optimization is a process to achieve a feasible solution to any problem. To define an optimization problem, we must design the objective or cost function to be minimized or maximized. In the constrained optimization approach, we also examine the constraints of the system. We can define a constraint as a limitation that is dictated by the variables in the problem to prevent physical impossibilities. The general form of a constrained optimization problem can be presented as follows:

$$\min f(t_1, t_2, \dots, t_n) \quad (1a)$$

$$\text{s.t. } g_i(t_1, t_2, \dots, t_n) = a_i, \quad \forall i \in (1, 2, \dots, p) \quad (1b)$$

$$h_j(t_1, t_2, \dots, t_n) \leq b_j, \quad \forall j \in (1, 2, \dots, q). \quad (1c)$$

where the objective function to be minimized is given in (1a), the equality constraints are given in (1b), and the inequality constraints are given in (1c).

When we consider such constrained optimization problems, we may define the optimality conditions so-called Karush-Kuhn-Tucker (KKT) conditions. These conditions define a set of conditions to have an optimal solution for the problem. The first condition indicates that all constraints defined in the problem must be satisfied. The second condition guarantees that the cost will not be improved at the optimal solution. The third condition implies that the Lagrange multipliers and its corresponding constraint must be zero. The last condition mentions that the optimal Lagrange multipliers must be non-negative.

III. TRAFFIC FLOW STRUCTURES

The traffic flow can be managed in two different structures in roads: One-way traffic and two-way traffic [1]. One-way traffic allows vehicles to move in only one predefined direction on the particular roadway. However, it is more common that vehicles can travel in both directions at the same time on the same roadway. Therefore, we have two-way traffic structures which are divided into lanes. The importance of each traffic flow structure becomes evident if it is implemented on the correct roadway. On the other hand, a roadway can also be designed as a mixture of both structures if necessary.

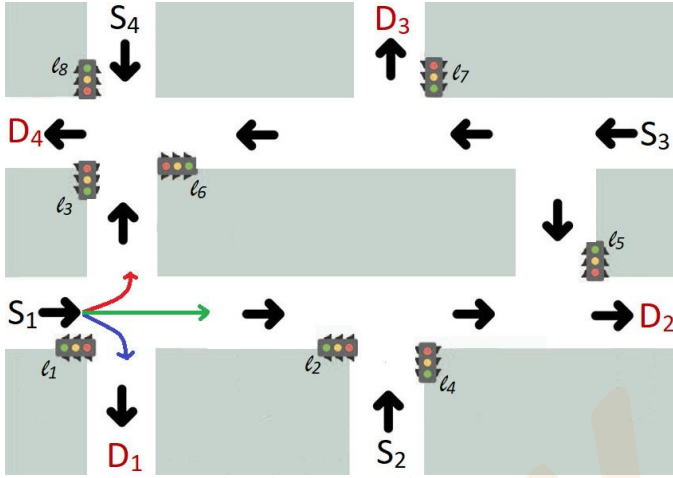


Fig. 1. A section of a city map with a one-way design.

Besides, we mainly consider strictly designated roadways to state our problem of interest. As can be seen from Fig. 1 the roadway is designed as one-way. The labels of S_i and D_j in Fig. 1 stand for possible Start and Destination points for a vehicle, respectively. For instance, a vehicle at S_1 can reach only D_1 , D_2 and D_4 in accordance with the permitted lane directions. Also there are traffic lights distributed around the map in order to manage traffic flow. Their locations can be organized as the needs of the lanes.

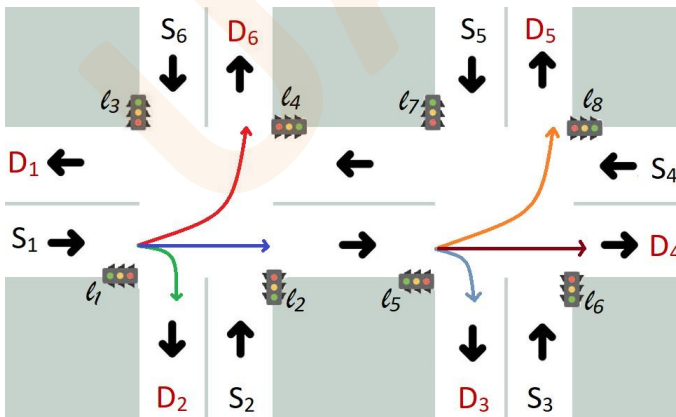


Fig. 2. A section of a city map with a two-way design.

The two-way structure can be more complicated than the one-way traffic flow designs since the junctions give more flexibility to vehicles for traveling to their destinations. In other words, a two-way structure gives more radius of action to vehicles. We illustrated a map of a two-way structured roadway in Figure 2. The labels of S_i and D_j in Fig. 2 stand for possible Start and Destination points for a vehicle respectively. We represented all the possible routes for a vehicle that starts from S_1 . For example, a vehicle can travel from S_1 to D_2 , D_3 , D_4 , D_5 and D_6 in the map illustrated in Fig. 2. The possible routes from another starting point to the destination can be determined in the same manner. We have again a number of traffic lights to control to optimize the number of vehicles in the two-way traffic. Although we can form several different roadway designs, the above two road structures can be considered as bases.

IV. PROBLEM DEFINITION

Most of the traffic light schedules are designated when they are first initialized. However, traffic light scheduling is essential to optimize the number of vehicles on roads. The management of the traffic light sequences by the vehicle density on roads allows us to minimize traffic congestion.

We utilize the constrained optimization technique to define our problem as follows:

$$\min \sum_{i=1}^n r_i - \mathbf{Lc}\mathbf{v} \quad (2a)$$

$$\text{s.t. } \mathbf{v} - \mathbf{v}^{\max} \leq 0 \quad (2b)$$

$$-\mathbf{v} \leq 0 \quad (2c)$$

$$\mathbf{Lc}\mathbf{v} - \sum_{i=1}^n r_i \leq 0 \quad (2d)$$

$$\mathbf{v}^{\min} - (\mathbf{r} - \mathbf{v}) \leq 0 \quad (2e)$$

where m defines the number of traffic lights, n defines the number of possible destination points on the map, $\mathbf{r}_i \in \mathbb{N}$ defines the number of vehicles at each route, $\mathbf{Lc} \in \mathbb{N}^{1 \times m}$ is the choice of the situation of traffic lights, $\mathbf{v} \in \mathbb{N}^m$ holds the number of vehicles passes when green light on, $\mathbf{v}^{\max} \in \mathbb{N}^m$ is the maximum vehicles can pass from each route i while l_j is green, \mathbf{v}^{\min} is the minimum vehicles should be waiting in queue at each route i while l_j is red.

We consider the cost function in (2a) to minimize the number of vehicles on the map in general. Since there are limited number of combinations for traffic lights to be green, the control center determines the L that minimizes the cost function most. The constraint (2b) determines the maximum number of vehicles that the road segment can hold. The constraint (2c) and (2d) impose the feasibility on the number of vehicles on the map. The last constraint indicates the minimum number of vehicles on each route.

V. RESULTS

In this section, we first review the necessary conditions for optimal solutions by using the KKT conditions. In addition,

we present a numerical example to show validity of the constrained optimization problem defined in Section IV.

A. Necessary Conditions for Optimal Solution

We first determine the necessary condition to find the optimal solution of the constrained optimization problem defined from (2a) to (2e). The Lagrangian can be presented as

$$\begin{aligned} \mathcal{L} = & \sum_1^n r - L * v + \lambda_1(v - V_{max}) + \lambda_2(-v) \\ & + \lambda_3\left(L * v - \sum_1^n r\right) + \lambda_4(V_{min} - r + v) \end{aligned} \quad (3)$$

The KKT conditions can be written as

$$\begin{aligned} -L - \lambda_1 - \lambda_2 + \lambda_3 L + \lambda_4 &= 0 \\ \lambda_1(v - V_{max}) &= 0 \\ \lambda_2(-v) &= 0 \\ \lambda_3\left(L * v - \sum_1^n r\right) &= 0 \\ \lambda_4(V_{min} - v) &= 0 \\ v - V_{max} &\leq 0 \\ -v &\leq 0 \\ \left(L * v - \sum_1^n r\right) &\leq 0 \\ V_{min} - v &\leq 0 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4, v &\geq 0. \end{aligned} \quad (4)$$

From the first condition in KKT conditions, we get $L(\lambda_3 - 1) - \lambda_1 - \lambda_2 + \lambda_4 = 0$. This leads to two cases:

$$\lambda_3 = 1 \text{ and } \lambda_4 - \lambda_1 - \lambda_2 = 0. \quad (5)$$

When we consider $\lambda_2 = 0$ from the third condition in (4), $\lambda_4 = \lambda_1$ is assured from (5). The complementarity conditions that are corresponding to λ_1 and λ_4 shows four cases:

Case 1: $v - V_{max} < 0$ leads to $v < V_{max}$ and $\lambda_1 = 0$. Thus $\lambda_4 = 0$. Then we conclude that $V_{min} + V_{max} < r$. This result can lead one of the optimal solutions.

Case 2: $v - V_{max} = 0$ leads to $v = V_{max}$. However, we may have two sub-cases for the lagrange multiplier λ_1 such $\lambda_1 = 0$ and $\lambda_1 > 0$. We have considered the former one in above case. We now consider $\lambda_1 > 0$. Thus $\lambda_4 > 0$. Then we get $V_{min} + V_{max} = r$ which is valid and can lead one of the optimal solutions.

Case 3: $V_{min} - r + v < 0$ leads directly the same conclusion in case 1.

Case 4: $V_{min} - r + v = 0$ leads directly the same conclusion in case 2.

Above conditions indicates the necessary conditions to have

B. A Numerical Example

On the other hand, we present a numerical example by considering the map illustrated in Fig. 1. We have utilized an Open-Source library CVX and Optimization Toolbox in Matlab to solve our constrained optimization problem given from (2a) to (2e).

There are eight routes defined in Fig. 1. The number of vehicles on each route given as

$$r = [30, 45, 85, 33, 27, 92, 25, 49]^T.$$

Therefore, there are currently 386 vehicles on road in total. The maximum and minimum capacities for each route are given as

$$\begin{aligned} v_{max} &= [0, 50, 50, 50, 50, 50, 0, 50]^T, \\ v_{min} &= [0, 15, 15, 15, 15, 15, 0, 15]^T. \end{aligned}$$

The choice of the situation of traffic lights \mathbf{L}_c can be chosen as one of the rows determined in the following matrix,

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Above matrix shows that there are nine combinations for eight traffic lights in this map. The value 1 in matrix \mathbf{L} indicates the green light while 0 is the red light.

The choice of \mathbf{L}_c changes the resultant cost in problem defined (2) immediately. Therefore, we need to assess all the possible traffic light configuration defined in \mathbf{L} to find the optimal solution for \mathbf{v} . The following table represents the optimal \mathbf{v} for each traffic light configuration and the cost in accordance with the optimal \mathbf{v} .

TABLE I

Choice of \mathbf{L} (row #)	optimal \mathbf{v}	cost
1	$[0, 30, 50, 0, 0, 0, 0, 0]^T$	306
2	$[0, 0, 50, 18, 0, 0, 0, 0]^T$	318
3	$[0, 0, 10, 0, 12, 0, 0, 0]^T$	324
4	$[0, 30, 0, 0, 0, 50, 0, 0]^T$	306
5	$[0, 0, 0, 18, 0, 50, 0, 0]^T$	318
6	$[0, 0, 0, 0, 12, 50, 0, 0]^T$	324
7	$[0, 30, 0, 0, 0, 0, 0, 34]^T$	322
8	$[0, 0, 0, 18, 0, 0, 0, 34]^T$	334
9	$[0, 0, 0, 0, 12, 0, 0, 34]^T$	340

optimal solutions of our problem of interest.

Above results represent that there are two different choices of traffic light configuration makes the cost minimum in map illustrated in Fig. 1. Therefore, either of the possible greenlight sequence can be chosen for minimizing the number of vehicles on the road.

The following matrix \mathbf{L} is formed in accordance with the possible traffic light configurations for the map given in Fig. 2.

$$\mathbf{L}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

There are sixteen possible actions for the traffic lights to turn green at the same time. We can directly implement one choice of configuration as \mathbf{L}_c in our problem definition. Then we can find an optimal solution by taking into account the current road conditions for map in Fig. 2.

VI. CONCLUSION

In this paper, we study a constrained optimization technique to minimize the traffic congestion on the roads. We review two types of traffic flow structures designed in one-way and two-way fashions. These two structures can be utilized to form the urban roadways. Therefore, roads may be designed to be more complex and intertwined which can cause traffic congestion. We propose a mathematical problem definition to minimize the number of vehicles on each route, regardless of the road design. We mainly try to supervise the traffic light configurations on routes. We have defined the necessary conditions to have the optimal solution of our problem. We have also provided a numerical example to validate our findings.

REFERENCES

- [1] Md Aftabuzzaman. "Measuring traffic congestion-a critical review". In: *30th Australasian transport research forum*. Vol. 1. ETM GROUP London UK. 2007.
- [2] Brian D Taylor. "Rethinking traffic congestion". In: *Access Magazine* 1.21 (2002), pp. 8–16.
- [3] Phil Lasley. *2019 Urban mobility report*. 2019.
- [4] Richard Arnott and Kenneth Small. "The economics of traffic congestion". In: *American scientist* 82.5 (1994), pp. 446–455.
- [5] Robin Lindsney* and Erik Verhoef*. "Traffic congestion and congestion pricing". In: *Handbook of transport systems and traffic control*. Emerald Group Publishing Limited, 2001, pp. 77–105.
- [6] Daniel Stokols et al. "Traffic congestion, type A behavior, and stress." In: *Journal of Applied Psychology* 63.4 (1978), p. 467.
- [7] Robert L Bertini. "You are the traffic jam: an examination of congestion measures". In: *The 85th annual meeting of transportation research board*. 2006, p. 115.
- [8] Timo Lajunen, Dianne Parker, and Heikki Summala. "Does traffic congestion increase driver aggression?" In: *Transportation research part F: traffic psychology and behaviour* 2.4 (1999), pp. 225–236.
- [9] Marlon G Boarnet, Eugene Jae Kim, and Emily Parkany. "Measuring traffic congestion". In: *Transportation Research Record* 1634.1 (1998), pp. 93–99.
- [10] Matthias Sweet. "Does traffic congestion slow the economy?" In: *Journal of Planning Literature* 26.4 (2011), pp.391–400

Model Training and Real World Analysis Using Health Data with Federated Learning

Abdülbaki Demir
Computer Engineering
Marmara University Faculty of
Technology
Istanbul, Türkiye
0009-0002-7603-9422

Ahmet Yasir Kulaksız
Computer Engineering
Marmara University Faculty of
Technology
Istanbul, Türkiye
0009-0007-0886-1795

Büşra Büyüktanır
Computer Engineering
Marmara University Faculty of
Technology
Istanbul, Türkiye
0000-0003-2571-4029

Gözde Karataş Baydoğmuş
Computer Engineering
Marmara University Faculty of
Technology
Istanbul, Türkiye
0000-0003-2303-9410

Kazım Yıldız
Computer Engineering
Marmara University Faculty of
Technology
Istanbul, Türkiye
0000-0001-6999-1410

Abstract—Machine learning has emerged in many fields thanks to its ability to extract meaningful information from large data sets, solve complex problems, and make predictions about the future. In healthcare, machine learning is used in various applications such as disease diagnosis, treatment planning, patient monitoring, and personalized healthcare. However, for these applications to be successful, large size of data are needed. Concerns about the privacy and confidentiality of health data make it difficult to collect this data on centralized servers and limit the effectiveness of data-driven models. Because of the need to protect privacy, collecting and analyzing individual patient data in a centralized system becomes a major challenge. Federated learning is a machine learning approach that allows data to be processed locally without being collected on a central server, and only model parameters are sent to the central server. This method provides secure model training by reducing network traffic while maintaining the confidentiality of the data. In this paper, we focus on diabetes prediction using the duCBA method developed for federated learning architecture. For model training, a dataset called Diabetes Health Indicators on the Kaggle platform was used. To manage the data flow and perform model training securely, three different environments were used: a central server, a Flask-based API server, and clients. Clients provide secure data privacy by performing distributed model training on their local data without sharing the data. The API and the central server are run in the Google Cloud Platform environment. The API server collects local model updates from clients and sends them to the central server using WebSocket when it has enough models. After the central server performs the model merge, it sends the updated global model to the API server. The API server then distributes this model to the clients. With this structure, clients participate in a centralized global model training process through the API without directly accessing the central server, and communication security is ensured through the WebSocket protocol. As a result of the tests conducted in the prototype application, the accuracy value of the model was calculated as 70%. These results support the wide applicability of federated learning in the healthcare domain.

Keywords— Federated Learning, Relational Classification, duCBA, Data Privacy, Diabetes Prediction, Machine Learning

I. INTRODUCTION

The increasing use of IoT devices with the developing technology has led to an increase in data generation[1]. Since these devices are used in many different areas, they offer various data collection opportunities. This data is used in many

areas from analyzing the user's behavior to monitoring environmental conditions [2]. While using this data, systems are created by creating prediction models with machine learning algorithms [3].

Machine learning is a type of artificial intelligence that can learn from data and identify complex patterns and relationships, make decisions on previously unseen data, or predict future outcomes without the need for programming [4]. For the model to perform effectively, large size of data need to be analyzed. However, increasing the size of data increases both the hardware capacity and the training time of the model [5]. In addition, the necessity of transmitting data to a single system during the training process brings problems such as storage and security. These issues become even more critical, especially in areas with large size of sensitive data, such as the healthcare industry. Therefore, secure storage and protection of health data is of great importance [6-7]. There are legal regulations in many countries regarding the collection and use of this data, and technological and legal solutions are being developed to protect the privacy of personal data [8]. However, these laws for the protection of personal data have created some difficulties for machine learning projects. In particular, restrictions on the sharing of personal data can lead to the difficulty of creating the data sets required for machine learning. This is an important problem for machine learning.

Federated learning is a different approach from traditional machine learning methods. In this method, the data to be used for model training is not collected on a central server. Instead, models are trained independently in local environments where the model will be used and data is generated. These trained independent models are then combined in another environment. Thus, a global model is created [9]. The most important advantages of this structure of federated learning are reducing the training time of the trained model, ensuring data privacy and reducing the bandwidth used in communication between the centralized structure and the client. In the traditional model training process, the training time increases as the training data increases. However, in federated learning, since the data is trained by dividing it into small pieces, the training time is reduced. In the traditional model training process, training data must be sent to a central server and this data must be stored in a centralized structure. In the federated learning method, since the models trained on local devices are sent instead of sending training data, it provides a solution to both the storage problem in the centralized structure and the

bandwidth problem that occurs during data transmission [10]. In addition, since the federated learning method does not send data to the centralized structure, the security of critical information on local devices is also ensured [11].

Taking all this into account, the advantages of federated learning are the following:

- **Data Privacy:** During the model training process, the data collected on local devices is trained without being sent to a central server. In this way, model training is provided without sharing the data.
- **Reduced Training Time:** Instead of processing large sizes at once, smaller data can be processed and combined across multiple devices, reducing training time.
- **Bandwidth Efficiency:** Sending trained models to the central server instead of data to be trained saves bandwidth.
- **Solves Storage Problems:** Since the data is not collected on a central server, the storage space requirement is reduced.
- **Compliance with Legal Regulations:** There are various legal regulations regarding the sharing of sensitive personal data. Federated learning protects user privacy as it processes data without sending it to a centralized entity and provides an effective solution to these regulatory challenges.

Federated learning architectures are used in many areas from education to health, security and finance. Especially in terms of protecting patient privacy, researches in the field of health should be important and recent. In this research, an application on diabetes prediction is developed using federated learning architecture. Diabetes is an increasing health problem worldwide and its early diagnosis is of great importance [12]. However, the privacy and security of medical data is a major obstacle to the application of traditional machine learning approaches. With federated learning, patient data is processed on local devices without transferring it to a central server, thus protecting patient privacy.

The next sections of the research examine the following topics. Section 2 reviews the work done for federated learning. Section 3 details the implementation techniques developed. Section 4 presents the results from the implementation. Section 5 discusses the overall conclusions of the research.

II. RELATED WORKS

In this section, the federated learning architecture and recent researches in the field of health within this architecture are presented.

Introduced in 2016 by Google to protect data privacy, the federated learning architecture has many advantages. Data privacy is ensured during model training and model training time is shortened. Since the model circulates between the clients and the server instead of data, bandwidth and communication latency problems are eliminated. In addition, storage costs are reduced since the model is collected on the central server instead of data. In addition to all these advantages of federated learning, there are some important problems that can be categorized according to four titles. These are communication overhead, system heterogeneity, statistical heterogeneity and privacy concerns.

Communication overhead is the difficulty in managing the model submission process as the number of local devices increases [13]. System heterogeneity is the difference in performance due to hardware and software differences between the devices on which model training is performed [14]. Statistical heterogeneity is caused by differences in the distribution of the data collected by the devices. This may cause the model to perform well on some devices and poorly on others [15]. In the federated learning method, although the confidentiality of important information is ensured by not sending any data to the central server for model training and only sending locally trained models, there is still a concern of accessing critical information by using the parameters of the models sent. This situation continues to raise privacy concerns. Algorithms have been developed to resolve privacy concerns [14-16].

Many methods for federated learning architecture have been developed in the literature. The Federated Gradient Information Aggregation (FedGiA) algorithm combines gradient aggregation and model updating in a hybrid approach to solve the problems of communication overhead and statistical heterogeneity. In this way, it aims to reduce communication costs and increase model accuracy [17]. Federated Asynchronous Update (FedAvu) method solves the system heterogeneity problem with asynchronous model updates, allowing participants to work at their own pace [18]. The Fair Federated Learning (FairFed) algorithm solves the statistical heterogeneity problem by focusing on providing balanced model performance across different groups [19]. The privacy-oriented Robustness of Secure Federated Learning (RoFL) method aims to ensure data privacy and security through secure gradient aggregation and cryptographic techniques that minimize the impact of malicious participants [20]. Another federated learning architecture that will be used in this research is the Data Unaware Classification Based on Association Rules (duCBA) method [21]. duCBA is a method used to combine models trained with the Classification Based on Association Rules (CBA) [22] algorithm on end devices. In this process, the support and confidence values of each rule are recalculated and the rules are ranked. When rules with the same label come from different models, the extent to which each rule is present in the dataset is checked and the support and confidence values are updated accordingly. For rules marked with different labels but with the same content, the frequency of occurrence is first compared; the rule with the higher support value is selected and included in the list. For equal confidence values, a ranking is made based on the support value. If the support values are the same, the first rule in the list is placed higher. The list of rules obtained at the end of all these processes forms the basis of the final model. duCBA algorithm was developed to ensure data privacy based on the relational classification approach. The main purpose of this algorithm is to train models with data locally on client devices instead of a centralized data collection and send only trained models to the server. This minimizes network traffic and data transfer while maintaining data confidentiality.

Federated learning is emerging as a powerful method to provide privacy and security requirements when developing machine learning models on medical data. This method protects data privacy by processing data on local devices, making it unnecessary to move patient data to a central server. Evaluations show that this privacy and security-oriented structure of federated learning is applicable to medical data and makes significant contributions to protecting patient

privacy when developing distributed data processing models [6-8].

Choudhury et al. (2019) [23] state that federated learning models supported by differential privacy enable local processing of sensitive health data without collecting it in centralized systems, thus significantly reducing the risk of patient data leakage. Thanks to this method, it becomes possible to analyze health data securely while protecting the privacy of individual data. Similarly, Horvath et al. (2023) [24] investigated the effectiveness of differential privacy and federated learning on the MIMIC-III dataset. In this research, it was shown that successful models can be built while preserving the confidentiality of patient data, and it was stated that decentralized learning processes enable medical data to be shared securely between different institutions. In Durga et al. (2021) [25], the Healthchain system, in which the federated learning model is integrated with blockchain-based security measures, is discussed. It is stated that this system enables secure data processing and data sharing without the need for a centralized intermediary. Thus, federated learning is considered to be the key to training more effective and reliable machine learning models on medical data while prioritizing privacy and security.

In a research by Chen et al. (2023) [26] on the comparison of machine learning algorithms and feature visualization analysis, the accuracy performance of different algorithms in diabetes prediction was examined and it was shown that certain algorithms were more successful in risk prediction. Similarly, in another research by Hama Saeed (2023) [27] for diabetes classification, the accuracy of machine learning algorithms was improved by using up-sampling technique to balance imbalanced data sets in type 2 diabetes classification. These researches show that machine learning algorithms offer powerful tools for diabetes risk prediction and classification, and techniques such as data balancing can further improve performance.

These researches show that federated learning is an important application area in healthcare and an active research process continues in this field. In particular, by seeking solutions to critical problems such as data privacy, security and system efficiency, federated learning enables powerful and effective machine learning models to be trained on medical data. In this context, the existing literature demonstrates the potential of federated learning to preserve privacy and improve efficiency in medical data. This research aims to provide a new contribution to the literature on the applications of federated learning methods in the healthcare domain.

III. METHOD

In this section of the research, the dataset used for model training, the technologies and methods used for the developed application are explained in detail.

A. Dataset

In conducting this research, the "Diabetes Health Indicators Dataset" [28] dataset available on the Kaggle platform was used to train and evaluate the models. The dataset is based on the Behavioral Risk Factor Surveillance System (BRFSS) survey data conducted by the US Centers for Disease Control and Prevention (CDC). This dataset contains information on diabetes and other factors that may affect individuals' general health. The dataset is balanced within both

classes and consists of 70,692 rows and 22 columns. The dependent variable "Diabetes_binary" column provides binary information about the respondent's diabetes status. The other 21 independent variables cover various factors representing individuals' lifestyles, health behaviors and general health status. These variables include high blood pressure, cholesterol level, body mass index (BMI), physical activity level, smoking, history of heart disease, general health status and income levels. Python programming language was used for data set analysis and visualization. In addition, data preprocessing and evaluation stages were carried out using Pandas, Scikit-learn, Numpy Seaborn and Matplotlib libraries.

Fig.1. Correlation of Independent Variables with the Dependent Variable

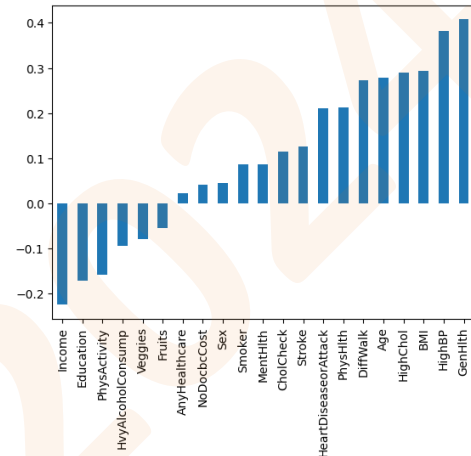


Fig.1 shows the independent variables in the data set and their correlation graph on diabetes. According to this graph, the effect of independent variables on diabetes can be positive or negative. Positive values increase the risk of diabetes, while negative values decrease the risk. According to the Fig. 1, 'GenHlth' and 'HighBP' are strongly associated with diabetes, which may increase the risk of diabetes. On the other hand, variables such as 'Income' are negatively associated with diabetes risk, suggesting that a higher income level may reduce the risk of diabetes.

B. Prototype Application

Three different environments were used in the prototype application developed within the scope of the research.

- Client
- API
- Central Server

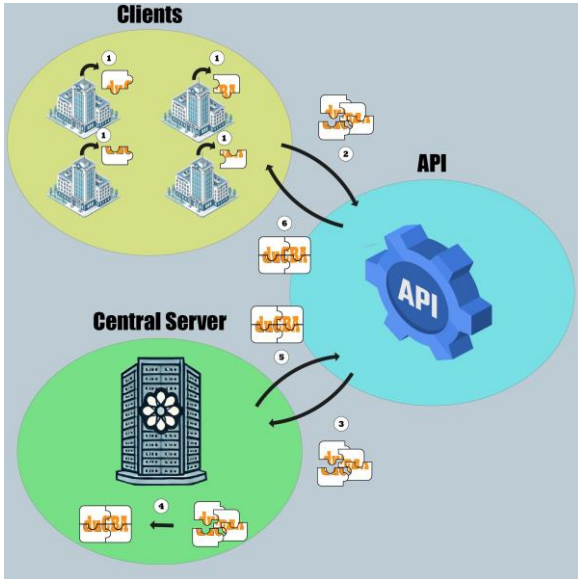


Fig.2. duCBA Workflow Diagram

In Fig. 2 shows the working principle of the prototype application. In the first step, local clients train models on their own data using the CBA method and send these models to the API server. The API server collects the local models and forwards them to the central server. The central server combines the incoming models with the duCBA method to create an updated model. This updated model is sent back to the API server. The API server distributes this updated model to the clients, enabling each client to use a more robust and global model. Fig. 2 shows this process as 1 round. Due to the structure of federated learning, the tour continues continuously and models trained with more data are trained. This process provides an approach that aims to learn on local data and training a global model in this way, while maintaining the confidentiality of the data.

1) Client

For the clients, 25 identical virtual machines were created on Google Cloud Platform. e2-medium machine type using x86/64 architecture was used. Python version 3.8.10 and numpy, pandas, scikit-learn, pickle, pyfim, requests, pyarc libraries were used for model training. During the model training process, the model was trained by randomly selecting data between 1000-3000 from the data set. CBA algorithm was used for model training. The support and confidence values we will use for CBA are set as 0.2 and 0.5 respectively. These values are pre-tested and approved values [21]. The models created and trained with these settings are sent to the API in json format by HTTP request. The API sends these models to the central server for merging. The models merged on the central server are sent to the clients by API and the clients continue their training with the global model.

2) API

For the API, e2-medium machine type was used on Google Cloud Platform. Python version 3.8.10 and flask, numpy, pandas, scikit-learn, pickle, pyfim, websockets, requests, pyarc libraries were used. Requests to the API were received using port 5000. In the prototype application tests, the models coming from 25 clients were collected 10 times with the API

and sent to the central server. These models, which were sent to the central server by websockets, were received from the central server and distributed to the clients by API.

3) Central Server

For the central server, e2-medium machine type was used on Google Cloud Platform. Python version 3.8.10 and numpy, pandas, scikit-learn, pickle, pyfim, websockets, requests, pyarc libraries were used. Websocket connection was realized using port 8000. The duCBA algorithm was used to combine the models coming from the API. The algorithm aims to merge local models trained with the CBA method on each client on the server. During the merging process, the rules obtained from each client are thoroughly evaluated and merged by recalculating the support and confidence values. This evaluation process is performed in order to achieve an optimal result when similar or identical rules come from different clients. During merging, the support and confidence values are updated by checking how many rules marked with the same label but coming from different clients are present in the dataset. Thus, the validity of each rule for the whole system is calculated more accurately. The frequency of occurrence (support value) is compared between rules with the same content but marked with different labels, and the rule with the higher support value is selected and added to the final list. After the update, the rules are sorted according to the new confidence values and the support value is taken into account in case of equal confidence values. If the support values are the same, the first rule in the list is ranked higher.

This extensive evaluation and ranking process results in a unified model to which all rules from clients contribute. The models created by the clients with the API on the central server are merged with the duCBA algorithm as mentioned before and the global model created is sent back to the clients by API.

IV. RESULTS AND ANALYSIS

In this section, the results of the analysis of the data obtained within the scope of the research are presented in detail. The results include the data related to the hypotheses and research questions determined in accordance with the purpose of the research.

TABLE I. GLOBAL MODEL ACCURACY PER ITERATION

Iteration	Accuracy
1	0.71
2	0.70
3	0.69
4	0.70
5	0.70
6	0.70
7	0.69
8	0.70
9	0.69
10	0.70

Accuracy values of the global models combined in a total of 10 iterations for model training of 25 clients during the

testing process of the prototype application. It is shown in Table 1.

Table 1 shows that these values range between 0.69 and 0.71. In general, the accuracy of the model remained stable around 0.69 and 0.70, with small deviations throughout the iterations. This suggests that the model has a consistent performance but without significant improvement. To improve the performance of the model, methods such as hyperparameter adjustments, increasing the dataset size or early stopping can be tried.

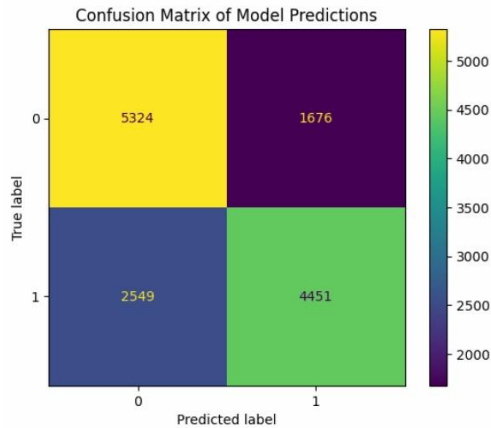


Fig.3. Global Model Confusion Matrix

Fig. 3 shows the confusion matrix created by predicting the global model obtained at the end of 10 iterations combined with the duCBA algorithm on 14000 test data. With the data in this matrix, precision, recall and f1-score values were calculated. These data are shown in table 2.

TABLE 2. DUCBA CLASSIFICATION REPORT

	Precision	Recall	F1-Score	Support
0	0.68	0.76	0.72	7000
1	0.73	0.64	0.68	7000
Accuracy			0.70	14000
Macro Avg.	0.70	0.70	0.70	14000
Weighted Avg.	0.70	0.70	0.70	14000

Table 2 shows the performance of the global model combined with duCBA on diabetes prediction. Values such as Precision, Recall and F1-score reveal how the model performs for each class. For class 0, Precision is 0.68, Recall is 0.76 and F1-score is 0.72. This shows that the model is relatively more sensitive in predicting class 0, but also produces some false positives. For class 1, Precision is 0.73, Recall is 0.64 and F1-score is 0.68, indicating that the model is more accurate in predicting class 1, but less successful than class 0 in finding samples that are actually class 1. The overall accuracy of the global model is given as 70%. This accuracy rate means that the model shows an overall balanced performance. The macro avg and weighted avg values are also similar for both classes at 0.70, indicating that there is no significant performance difference between the classes.

The results obtained throughout the research show that the duCBA algorithm can be successfully applied on health data in a federated learning approach.

V. CONCLUSION

In this research, the performance of a federated learning based model on diabetes prediction is analyzed. As a result of the tests, promising results were obtained in model performance by using the federated learning architecture and duCBA algorithm. These results show that the combination of federated learning and duCBA contributes to the development of an effective prediction model, especially on sensitive health data. In the future, the applicability of this model for predicting different diseases and in other health domains can be explored. Furthermore, it is recommended to integrate advanced algorithms for security and performance improvements and to conduct more extensive testing in real world applications.

REFERENCES

- [1] Sasaki, Y., A survey on IoT big data analytic systems: Current and future. *IEEE Internet of Things Journal*, 9(2), 1024–1036, 2022.
- [2] John Dian, F., Vahidnia, R., & Rahmati, A., Wearables and the Internet of Things (IoT), applications opportunities and challenges: A survey, *IEEE Access*, 8, 69200–69211, 2020.
- [3] Bian, J., Arafat, A. A., Xiong, H., Li, J., Li, L., Chen, H., Wang, J., Dou, D., & Guo, Z., Machine learning in real-time Internet of Things (IoT) systems: A survey. *IEEE Internet of Things Journal*, 9(11), 8364–8386, 2022.
- [4] Telikani, A., Tahmassebi, A., Banzhaf, W., & Gandomi, A. H., Evolutionary machine learning: A survey. *ACM Computing Surveys*, 54(8), Article 161, 2021.
- [5] Paleyes, A., Urma, R.-G., & Lawrence, N. D., Challenges in deploying machine learning: A survey of case studies. *ACM Computing Surveys*, 55(6), Article 114, 1–29, 2022
- [6] Choudhury, O., Gkoullas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G., & Das, A., Differential privacy-enabled federated learning for sensitive health data, 2019
- [7] Zhou, S., & Li, G. Y., FedGiA: An efficient hybrid algorithm for federated learning, *IEEE Transactions on Signal Processing*, 71, 1493-1508, 2023.
- [8] Sun, Y., Li, H., Shen, Y., Xie, J., Zhao, Y., Gao, X., & Si, N., An Asynchronous Federated Learning Algorithm Based on a Backup Update of Model Version Parameters, 2023 3rd International Conference on Electrical Engineering and Control Science (IC2ECS), 1538-1544, 2023.
- [9] Raj, A., Sharma, V., & Shanu, A. K., Comparative analysis of security and privacy techniques for federated learning in IoT-based devices, In 2022 3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM), 1-5, 2022.
- [10] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V., Federated Learning: Challenges Methods and Future Directions, *IEEE Signal Processing Magazine*, 37(3), 50-60, 1, May, 2020.
- [11] Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W., A survey on federated learning: Challenges and applications, *International Journal of Machine Learning and Cybernetics*, 14(2), 513-535, 2023.
- [12] Ali, M. K., Pearson-Stuttard, J., Selvin, E., & Gregg, E. W., Interpreting global trends in type 2 diabetes complications and mortality. *Diabetologia*, 65(1), 3–13, 2022.
- [13] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V., Federated Learning: Challenges Methods and Future Directions, *IEEE Signal Processing Magazine*, 37(3), 50-60, 1, May, 2020.
- [14] Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W., A survey on federated learning: Challenges and applications, *International Journal of Machine Learning and Cybernetics*, 14(2), 513-535, 2023.

- [15] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y., A survey on federated learning, *Knowledge-Based Systems*, 216, 106775, 2021.
- [16] Kaur, H., Rani, V., Kumar, M., Sachdeva, M., Mittal, A., & Kumar, K., Federated learning: A comprehensive review of recent advances and applications, *Multimedia Tools and Applications*, 83(18), 54165-54188, 2024.
- [17] Zhou, S., & Li, G. Y., FedGiA: An efficient hybrid algorithm for federated learning, *IEEE Transactions on Signal Processing*, 71, 1493-1508, 2023.
- [18] Sun, Y., Li, H., Shen, Y., Xie, J., Zhao, Y., Gao, X., & Si, N., An Asynchronous Federated Learning Algorithm Based on a Backup Update of Model Version Parameters, *2023 3rd International Conference on Electrical Engineering and Control Science (IC2ECS)*, 1538-1544, 2023.
- [19] Ezzeldin, Y. H., Yan, S., He, C., Ferrara, E., & Avestimehr, A. S., Fairfed: Enabling group fairness in federated learning, In *Proceedings of the AAAI conference on artificial intelligence*, 37(6), 7494-7502, June, 2023.
- [20] Lycklama, H., Burkhalter, L., Viand, A., K uchler, N., & Hithnawi, A., RoFL: Robustness of Secure Federated Learning, *2023 IEEE Symposium on Security and Privacy (SP)*, 453-476, 2023.
- [21] B Y KATANIR, B., YILDIZ, K.,  LK , E. E., & B t kτανir, T., du-CBA: Data-agnostic and incremental classification-based association rules extraction architecture du-CBA: Veriden habersiz ve artirimli siniflandirmaya dayali birliktelik kurallari cikarma mimarisi. *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol.38, no.3, 1919-1929, 2023
- [22] Abouelmehdi, K., Beni-Hssane, A., Khaloufi, H., & Saadi, M., Big data security and privacy in healthcare: A review. *Procedia Computer Science*, 113, 73-80, 2017
- [23] Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G., & Das, A., Differential privacy-enabled federated learning for sensitive health data, *arXiv preprint arXiv:1910.02578*, 2019.
- [24] Horvath, A. N., Berchier, M., Nooralahzadeh, F., Allam, A., & Krauthammer, M., Exploratory Analysis of Federated Learning Methods with Differential Privacy on MIMIC-III, *arXiv preprint arXiv:2302.04208*, 2023.
- [25] Durga, R., & Poovammal, E., Federated learning model for healthchain system, In *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, 1-6, 2021.
- [26] Chen, S., Comparison of machine learning algorithms and feature visualization analysis for diabetes risk prediction, *Journal of Physics: Conference Series*, 2646(1), 012013, 2023
- [27] Hama Saeed, M. A., Diabetes type 2 classification using machine learning algorithms with up-sampling technique, *Journal of Electrical Systems and Information Technology*, 10(1), 8, 2023.
- [28] Centers for Disease Control and Prevention (CDC), "Diabetes Health Indicators Dataset", 2024, Kaggle. [Online]. Available: <https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset/data>

Development of a Cost-Efficient FPGA Carrier Board for Digital Logic Design Courses in Electronics Engineering Curricula

Mehmet Fırat Aydın
Δ Laboratory, Dept. of Electronics Eng.
Gebze Technical University
Kocaeli, Türkiye
m.aydin2021@gtu.edu.tr

İhsan Çiçek
Δ Laboratory, Dept. of Electronics Eng.
Gebze Technical University
Kocaeli, Türkiye
ih sancicek@gtu.edu.tr

Abstract— Economic challenges faced by developing and underdeveloped countries often limit undergraduate electrical-electronics engineering students' access to hands-on learning in digital logic design platforms, potentially hindering the development of a skilled workforce. Practical experience is crucial for bridging the gap between theoretical knowledge and industry requirements, yet resource constraints often make it difficult to provide such opportunities. This paper introduces the design and initial prototype of a cost-efficient FPGA carrier board tailored for undergraduate digital design courses, while also being applicable for research and graduate-level courses. The board measures 12.75 x 12.75 cm and supports Sipeed GOWIN based FPGA system on modules, which are compatible with the available free and open-source toolchain for the design and implementation of custom logic circuits. With its projected low cost, this board is designed to enhance the accessibility and scalability of practical digital design education for students in economically disadvantaged regions.

Keywords—FPGA, logic design, boolean, digital design, engineering education, RISC-V, embedded systems

I. INTRODUCTION

Economic disparities significantly impact access to quality engineering education, particularly in developing and underdeveloped countries. In electrical-electronics engineering, where hands-on experience with tools and platforms is essential, the lack of accessible, affordable hardware presents a substantial barrier to student learning and skill acquisition. Specifically, practical exposure in digital logic design courses is critical for equipping students with the competencies needed to meet industry demands. However, the high cost of educational tools and limited access to cutting-edge platforms often prevent institutions in resource-constrained regions from providing adequate learning opportunities. This creates a substantial skills gap and limits the ability of students to transition effectively into the workforce.

Digital logic design, a foundational area in electrical and electronics engineering curricula, requires students to work with hardware platforms such as field-programmable gate arrays (FPGAs). Thanks to their reconfigurable nature, FPGAs offer versatility and flexibility in implementing and testing digital circuits, making them an indispensable platform for courses on hardware description languages, finite state machines, and computer architecture. Despite their importance, commercial FPGA boards remain prohibitively expensive for many educational institutions in economically disadvantaged regions around the globe. Democratization of access to digital design platforms such as FPGA development

boards or kits often fails due to imbalance between affordability and functionality. As a result, students are often restricted to theoretical learning, which undermines their ability to develop practical skills and creative problem-solving abilities. Hands-on learning in engineering education is vital for producing the highly skilled engineers required by the electronics industry. Integrating practical experience with theoretical coursework is essential for achieving this objective, as it enhances student engagement and promotes better knowledge retention.

Several efforts have been made to address these challenges by developing affordable alternatives to conventional commercial FPGA platforms. For example, educational boards such as Digilent Basys3 and Terasic DE10-nano have been widely adopted in undergraduate courses offered in developed countries [1], [2]. However, these solutions still carry price points that are often unattainable for institutions in developing or underdeveloped regions of the globe. Additionally, many commercial solutions rely on proprietary EDA software ecosystems, which can further increase costs through hidden licensing costs and limit accessibility. Research has consistently shown that reducing barriers to entry through cost-effective, open-source platforms can significantly improve educational outcomes and foster innovation [3].

This paper presents a novel solution to these challenges: a cost-efficient FPGA carrier board designed for undergraduate digital logic design courses. Unlike existing platforms, this board leverages the low-cost Sipeed Tang Nano FPGA modules and integrates a free and open-source toolchain for custom logic circuit design [4], [5]. Measuring 12.75 x 12.75 cm and anticipated to cost under \$30 per unit, this board offers a scalable and accessible alternative for institutions in economically constrained regions.

The remainder of the paper is organized as follows: Section II discusses the hardware architecture of the FPGA carrier board, highlighting its technical specifications and integration process. Section III explores the PCB design and hardware prototyping, while Section IV provides potential use cases of the proposed platform in education. The paper concludes in Section VI with a summary of contributions and directions for future work.

II. HARDWARE ARCHITECTURE

In this section, we outline the design choices and hardware architecture of the proposed FPGA carrier board. As shown in Fig. 1, the board offers multiple DC power supply input

options, enhancing its versatility for both laboratory and field use. It can be powered via USB-C for programming or through a DC barrel jack or header, allowing standalone operation in scenarios such as robotic competitions or off-laboratory activities.

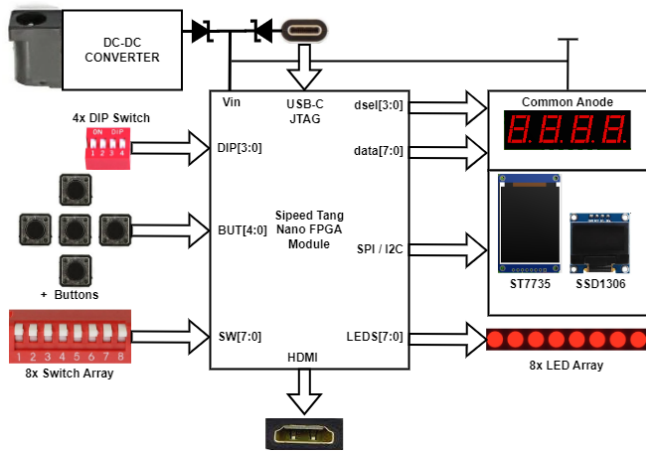


Fig. 1. System level hardware architecture of the FPGA carrier board.

To facilitate human-machine interaction, the board includes an array of 8 LEDs aligned with slide switches. These components support both input and output operations, making them ideal for combinational logic design experiments. Additionally, the board features a 4-digit multiplexed common anode seven-segment display module, enabling experiments with BCD number representation and digit multiplexing using FPGA resources as presented in Fig. 1. For advanced interfacing, the carrier board supports either an SPI-based ST7735 LCD or an I2C-compatible SSD1306 OLED display. These options are particularly useful for embedded soft-core CPU projects, such as implementing the RISC-V-based PicoRV32 processor on the FPGA module. Furthermore, all FPGA pins are routed to male pin headers, facilitating debugging and analysis using tools like logic analyzers or oscilloscopes.

The proposed FPGA carrier board serves as a cost-effective platform for developing, testing, and deploying complete design solutions. It is tailored to digital logic design courses and embedded system design applications, supporting experiments without requiring additional external components. The inclusion of sufficient pin headers also allows for the integration of custom add-on PCB modules to extend functionality. The peripherals and modules included for course-related experiments are discussed in the following sections.

A. Gowin GW1NR-9 FPGA Device

To balance manufacturing costs and design simplicity, we chose the low-cost Sipeed Tang Nano 9K module, shown in Fig. 2, as the core of our hardware design. Its compact form factor supports breadboard use and includes a USB-C connection for JTAG programming, debugging, and powering. The module also provides a full-size HDMI output and a dedicated socket for driving LCD modules up to 5 inches, enhancing its versatility for a range of applications. At the heart of the Tang Nano 9K module shown in Fig. 2 lies the GOWIN GW1NR-9 FPGA chip, which offers an excellent cost-to-performance ratio. This FPGA features 8640 logic

units (LUT4) and 6480 registers, complemented by 17,280 bits of ShadowSRAM (SSRAM) and 468 kilobits of Block SRAM (BSRAM) across 26 blocks [4]. With an onboard 27

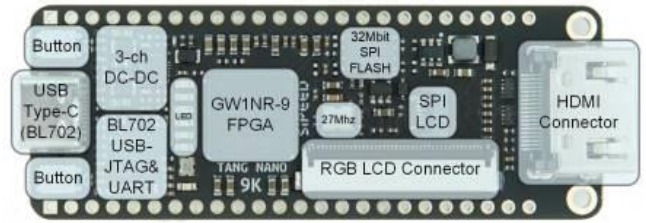


Fig. 2. Sipeed Tang Nano 9K FPGA board top view [4].

MHz reference clock and two PLLs, the FPGA is well-suited not only for basic designs but also for implementing full RISC-V soft-core CPUs such as PicoRV32 for use in embedded systems laboratory courses.

The Tang Nano 9K module in Fig. 2 includes display interface options, such as HDMI, RGB screen, and SPI screen interfaces, making it ideal for multimedia and embedded applications. It also features 32 Megabits of SPI flash and supports up to 64 Megabits of SDR SDRAM, offering ample memory for complex designs. Debugging is streamlined with an onboard BL702 chip, which supports USB-JTAG and USB-UART functions for the GW1NR-9 FPGA [4]. Flexible I/O capabilities further enhance the board's utility, with driving strengths of 4mA, 8mA, 16mA, and 24mA. Each I/O pin includes independent Bus Keeper functionality, configurable pull-up/pull-down resistors, and Open Drain outputs. Additional connectivity is provided through a TF card slot and two 24-pin 2.54mm headers, enabling data logging for IoT and wireless sensor network applications. User interaction is supported by two programmable buttons and six onboard LEDs, as shown in Fig. 2. These features facilitate customization and experimentation, making the module a valuable tool for both educational and prototyping environments.

Compared to alternative FPGA modules, the Tang Nano 9K stands out for its superior balance of logic capacity, memory, and peripheral options at a low cost. With its HDMI high-speed output, support for OLED and LCD displays, and multiple power supply options, it meets the needs of modern digital design challenges. These features make it an excellent choice for undergraduate level digital logic courses, embedded system projects, and microprocessor design, ensuring cost-efficient, safe, and rapid development without compromising functionality.

B. Power Supply Design

The FPGA carrier board is equipped with two standalone DC input options: a 2.1mm DC barrel jack and a 5.08mm standard terminal block as shown in Fig. 4. To safeguard against potential short circuits, a Schottky diode-based logical OR configuration is implemented. These dedicated input options make the board suitable for off-laboratory use or standalone deployments in various applications. The power supply circuitry shown in Fig. 3 includes two LM2596-based DC-DC converters to generate the 3.3V and 5V DC voltages required by the onboard peripherals and the FPGA module [6]. Alternatively, the FPGA module can be powered directly via its USB-C port. However, to prevent a race condition on the supply rails—which could lead to damage to the FPGA

module—it is critical to avoid simultaneously connecting an external DC power supply and the USB-C power source. The

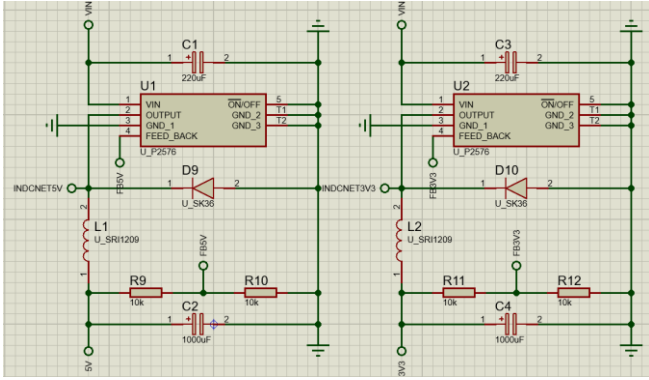


Fig. 3. LM2596 based FPGA carrier board power supply circuits.

board should operate using either the external DC input or the USB-C power, but not both at the same time. This flexible and robust power supply design ensures reliable operation for a wide range of use cases, whether in the lab or standalone environments.

C. General Purpose Input and Outputs

The FPGA carrier board is equipped with an ample number of GPIOs to support various input and output operations. It includes an array of eight LEDs and eight slide switches, complemented by four DIP switches, as shown in Fig. 1, providing a versatile platform for experimenting with digital logic designs. To enhance user interaction, the board also features five push buttons arranged in a joystick layout. This configuration offers flexibility for scenarios such as remote-control applications or custom user interfaces. All 48 pins of the Sipeed Tang Nano 9K FPGA module are routed to pin headers, allowing easy connection to external circuits. This design facilitates the expansion of the board's functionality, enabling users to integrate custom peripherals or debug and monitor FPGA pins using external tools like oscilloscopes or logic analyzers.

D. Display Interfaces

The FPGA carrier board includes a 4-digit, time-division multiplexed, common-anode seven-segment LED display, as shown in Fig. 5. This module is ideal for experimenting with decoders, counters, and BCD-encoded number displays, making it a valuable tool for understanding fundamental digital logic concepts.

For more advanced projects, such as those involving synthesizable soft processor cores, the board supports I2C- and SPI-capable display modules. These include both LCD and OLED options, offering flexibility for implementing graphical interfaces or data visualization tasks. Additionally, the board features an array of eight LEDs, each aligned with a corresponding slide switch, as shown in Fig. 5. This arrangement provides intuitive feedback and control, making it suitable for experiments with combinational and sequential logic designs.

III. PCB DESIGN AND HARDWARE PROTOTYPING

Following several design reviews and iterations, we developed a prototype PCB using Proteus 8 CAD software, as shown in Fig. 4. To minimize costs and ensure manufacturability under limited conditions, a two-layer stack-

up was selected, using FR4 material for its availability and durability. The target PCB dimensions were set at 12.7 x 12.7 cm, with careful component placement to ensure both easy routing and user accessibility. Embedded 3D models of components and display modules were utilized to perform dimensional fitness checks, enhancing design precision, as illustrated in Fig. 5.

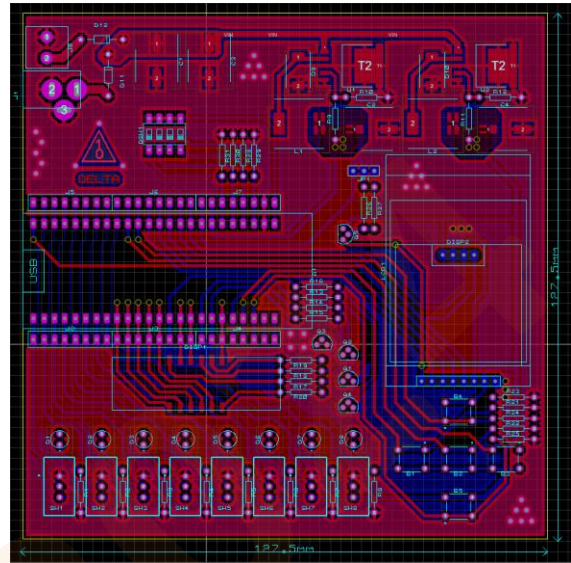


Fig. 4. PCB layout of the FPGA carrier board for prototype fabrication.

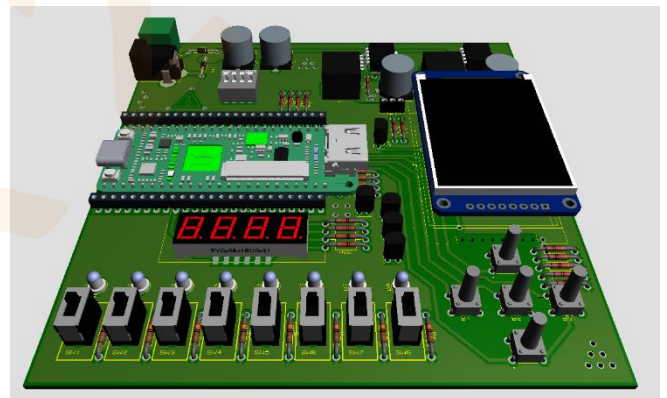


Fig. 5. 3D render view of the FPGA carrier board PCB with components.

After completing the layout, we conducted a Design Rule Check (DRC) and layout-versus-schematic verification to ensure design integrity. Gerber and drill files were then generated, and the PCB was fabricated using an in-house LPKF S100 PCB prototyping machine. Due to the limitations of this manufacturing method, which does not support via implementation, manual through-hole wires were used to replicate via functionality.

Once the board was carved, circuit nodes were tested for electrical connectivity using a multimeter in continuity mode. Following verification, electronic components were populated and soldered to prepare the board for initial testing, as shown in Fig. 6. The assembled PCB was cleaned with isopropyl alcohol to remove flux residues and ensure a clean surface. Before mounting the FPGA module, power supply connections were double-checked, and the correct voltages were verified at the corresponding power pins of the FPGA module. Initial tests confirmed the functionality of the FPGA carrier board design. The bare prototype PCB in Fig. 6,

however, lacks protective coating and is susceptible to oxidation over time.

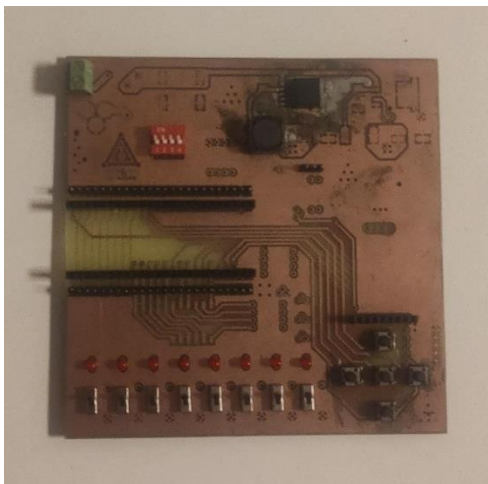


Fig. 6. First prototype of the FPGA carrier board PCB with component assembled and soldered.

Based on the successful validation of the initial design, a revised PCB is currently being developed for factory-level manufacturing. The new design will fit within a 10 x 10 cm footprint to reduce fabrication costs and mitigate potential customs-related issues.

IV. POTENTIAL USE CASES OF THE PROPOSED PLATFORM

In this section, we outline and explain the potential use cases of the proposed FPGA carrier board in a variety of academic scenarios and courses. The versatile features provided by the board make it a valuable tool for enhancing hands-on learning and experimentation across several engineering disciplines.

A. Digital Logic Design Laboratory Course

The FPGA carrier board is particularly suited for use in digital logic design laboratories, where students explore fundamental concepts such as combinational and sequential circuit design. With its integrated LED and switch arrays, the board provides a straightforward platform for implementing logic gates, multiplexers, decoders, counters, and finite state machines. The inclusion of a 4-digit seven-segment display allows students to visualize BCD outputs and perform experiments involving digit multiplexing. The GPIO accessibility ensures compatibility with custom circuits, encouraging students to develop creative solutions to design challenges.

B. Embedded System Design Laboratory Course

In embedded systems laboratories, the board's support for I2C- and SPI-capable display modules and onboard peripherals makes it ideal for implementing and testing embedded applications. Students can experiment with soft processor cores, such as the RISC-V-based PicoRV32, and interact with external devices like sensors, actuators, and memory modules. The board's USB-C and standalone power options ensure flexibility, whether the focus is on in-lab debugging or field deployments. Additionally, the versatile connectivity options and programming interfaces simplify integration with various embedded projects.

C. Unmanned Systems Laboratory Course

The proposed board can also play a critical role in unmanned systems laboratories, where students design and prototype control systems for drones, robots, and other autonomous platforms. The FPGA's high-speed processing capabilities and flexible GPIOs enable the development of sensor fusion algorithms, motor controllers, and communication interfaces. The board's compact design, combined with its standalone power options, makes it ideal for deploying in off-lab scenarios such as robotics competitions or outdoor experiments.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented the design and development of a cost-effective FPGA carrier board tailored to enhance hands-on learning in electronics engineering education. With features such as GPIOs, display modules, onboard peripherals, and flexible power options, the board offers a versatile platform for courses in digital logic design, embedded systems, and robotics. Its compact size and low manufacturing cost make it well-suited for academic use and extracurricular activities, including robotics competitions. The prototype has been validated through testing, confirming its functionality for educational purposes. Future work will focus on expanding the board's capabilities to support more advanced applications. Planned enhancements include wireless connectivity via ESP32 for IoT projects, sound generation for audio experiments, motor control for robotics, and ADC/DAC modules for analog signal processing. While these upgrades will broaden the board's functionality, maintaining affordability will remain a priority. Extended testing in diverse academic settings will guide future refinements, ensuring the board continues to serve as a cost-effective and adaptable educational tool.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Dr. Murat Ceylan for his invaluable support during the PCB prototyping phase of this study.

REFERENCES

- [1] "Basys 3 - Digilent Reference." Accessed: Nov. 2, 2024. [Online]. Available: <https://digilent.com/reference/programmable-logic/basys-3/start>
- [2] "Terasic - All FPGA Boards - Cyclone IV - DE0-Nano Development and Education Board." Accessed: Nov. 2, 2024. [Online]. Available: <https://www.terasic.com.tw/cgi-bin/page/archive.pl?Language=English&CategoryNo=139&No=593&PartNo=2>
- [3] B. NAJI, C. ABDELMOULA, K. ABBES, and M. MASMOUDI, "Design and test of a new development FPGA board for mobile robot research," *Turk. J. Electr. Eng. Comput. Sci.*, vol. 25, no. 2, pp. 1483–1494, Jan. 2017, doi: 10.3906/elk-1510-18.
- [4] "Tang Nano 9K - Sipeed Wiki." Accessed: Sep. 27, 2024. [Online]. Available: <https://wiki.sipeed.com/hardware/en/tang/Tang-Nano-9K/Nano-9K.html>
- [5] *YosysHQ/oss-cad-suite-build*. (Sep. 25, 2024). Shell. Yosys Headquarters. Accessed: Sep. 27, 2024. [Online]. Available: <https://github.com/YosysHQ/oss-cad-suite-build>
- [6] "LM2596 data sheet, product information and support | TI.com." Accessed: Sep. 27, 2024. [Online]. Available: <https://www.ti.com/product/LM2596>

Solving Linear Algebraic Equations by using Distributed Memory Programming

Oğuzhan Çiftçi

Dept. of Electrical & Electronics Engineering

Marmara University

Istanbul, Türkiye

0000-0002-9763-3677

Abstract—In this paper, we study solving of linear algebraic equations with the form of $Ax = b$ in a distributed manner by using distributed memory programming. We consider open-source message passing interface (MPI) approach to model our simulations. Also we review the sequential programming approach and make comparison with MPI on the performances under different network topologies.

Keywords—Multi-core Systems, Algebraic Equations

I. INTRODUCTION

Parallel programming techniques are widely used in computer simulations to enhance the computational performance. Nowadays, the computing systems are mostly designed as having multiple cores in their architectures and these cores can be used for parallel programming. One of the main benefits of having the multiple cores in computing systems is that these cores can be utilized individually at the same time. Therefore, programmers have the opportunity to use these cores a networked system. In the networked systems, each core processes their workloads and exchange their results with their connected neighbors [1-5].

There are numerous types of techniques to achieve parallel programming structure. One of the widely used techniques consider that each core processes their workload individually and has its own local memory in the computing system. Therefore, each core can be modelled as an individual processor and data in the local memory of the cores should be explicitly exchanged between the other cores. This structure is known as distributed memory systems. The Message Passing Interface (MPI) is a powerful tool to utilize the distributed memory programming.

On the other hand, most of the programmers do not spend enough time to benefit from the parallel programming tools [6]. In order to draw attention into parallel programming techniques, we suggest a use case for integrating these tools with the conventional multi-core computing systems.

Linear algebraic equations (LAEs) are very essential in order to analyze systems mathematically in many different fields such as Engineering, Economics, and Architecture. We usually use the following form to show the algebraic equations:

$$Ax = b \quad (1)$$

where $A \in \mathbb{R}^{m \times n}$ is the system matrix, $x \in \mathbb{R}^n$ is the solution vector, and $b \in \mathbb{R}^m$ is the observation vector [7].

There are numerous different algorithms suggested to solve such systems, however, they can be investigated in two major topics: Centralized algorithms and Distributed algorithms. Centralized algorithms generally consider the system altogether and try to manipulate the system matrix in order to make it invertible. On the other hand, distributed algorithms use a different approach. The main idea for the distributed algorithms is partitioning the system matrix into many parts and solve these parts individually by using different solvers. Then collect the solutions from each solver in accordance with a connection network between the solvers. Finally, solvers have a consensus on a solution for the overall system $Ax = b$ if there is a solution [7].

In this paper, we mainly focus on solving linear algebraic equations of the form (1) by using the distributed memory programming technique in C language. We use the MPI model which was designed to utilize the distributed memory architecture in multi-core computing systems.

This paper is organized as follows. In the following section, we provide mathematical preliminaries to form a network and introduce an algorithm to solve linear algebraic equation systems in a distributed manner. In the next section, we form numerical algebraic equation systems to be simulated in a multi-core computer. We introduced the pseudocodes for solving such equation systems both in sequential programming and distributed memory programming techniques. Section 5 compares the sequential programming and distributed memory programming simulation results. In section 6, we conclude and discuss the results provided in this paper.

II. MATHEMATICAL PRELIMINARY

In a multi-core computing system, we consider each core as an individual processor as mentioned previously. However, the data exchange relationship should be well-defined to have distributed memory structure at hand. Communication channels between the agents in a multi-core network can be expressed mathematically by using graph theory. We consider a directed graph $G = (V, E)$ as illustrated in Fig. 1, where $V = \{1, 2, \dots, m\}$ is the set of nodes and $E \subset V \times V$ is the set of directed edges between the nodes. We assume node i and node j are adjacent to each other if a communication link exists between them, i.e., $(i, j) \in E$. These nodes are also called *neighbors*. Furthermore, we define the set of nodes that

sends information to node i as in-neighbors of node i , i.e., $\mathbf{N}_i^+ = \{j \in \mathbf{V} : (j, i) \in \mathbf{E}\}$.

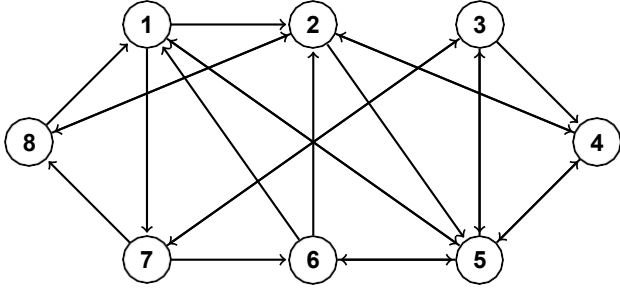


Fig. 1. A directed graph G .

The nature of the distributed algorithms for solving algebraic equations fits into the parallel programming structure. Especially, when we consider each core in the CPU as a node of a network, we can directly benefit from the parallel programming techniques in order to simulate the solution of the algebraic equations by using an appropriate algorithm. There are many different algorithms that can be utilized to solve the linear algebraic equation of the form of (1) distributively. However, we use a discrete-time distributed algorithm suggested in Mou et al., 2015 such that

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) - P_{ker(A_i)} \sum_{j \in \mathbf{N}_i^+(t)} w_{ij}(t) (\mathbf{x}_i(t) - \mathbf{x}_j(t)) \quad (2)$$

where $w_{ij}(t)$ is the weighting coefficient which was designed to be $\mathbf{N}_i^+(t)^{-1}$, $\forall i \in \mathbf{V}$, $P_{ker(A_i)}$ is the projection matrix onto the kernel of A_i , and $t \geq 1$.

III. PROBLEM DEFINITION

In this section, we define two different linear algebraic equation systems numerically. The first system has the following equations:

$$\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ 2x_1 + 3x_2 + x_3 &= 3 \\ x_1 - x_2 - 2x_3 &= -6. \end{aligned} \quad (3)$$

Above algebraic equation system has a unique solution of $\mathbf{x} = [-1, 1, 2]^T$. In order to solve a linear algebraic equation by using (2), we form a network with three nodes as illustrated in Fig. 2 and it satisfies all the conditions defined in Mou et al., 2015.

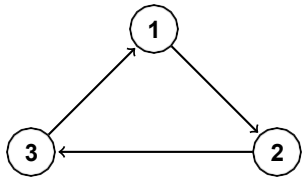


Fig. 2. A directed graph with three nodes.

As described in Mou et al., 2015, each node in the network should know a subset of the system (3). In addition, each node must begin its iteration in algorithm (2) with a solution to its equations.

Secondly, we design a system with eight equations of the form of (1) as follows:

$$\begin{bmatrix} 4 & -2 & 3 & 2 & 1 & 6 & 3 & 0 \\ 6 & 0 & 1 & -1 & -3 & 0 & -3 & -1 \\ -1 & -3 & 6 & -2 & 5 & -1 & -3 & 4 \\ 1 & 2 & -2 & -1 & -2 & -3 & 4 & -3 \\ 1 & 0 & 4 & 5 & -3 & -1 & 6 & -3 \\ 4 & -2 & 4 & -3 & 0 & -3 & 2 & 3 \\ 5 & -1 & 2 & -1 & 1 & 2 & -2 & 3 \\ -2 & 6 & -2 & -3 & -2 & 4 & 5 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 33 \\ 23 \\ 11 \\ 2 \\ 13 \\ 12 \\ 23 \\ 7 \end{bmatrix} \quad (4)$$

We suggest the network illustrated in Fig. 1 as the multi-core system to solve above equation system. Although it has infinitely many solutions, we can still use algorithm (2) to achieve consensus among the cores as a solution to overall system (4).

Furthermore, we simulate different randomly generated equation systems for simulation time comparison. Each system is guaranteed to have at least one solution and we utilize the network topologies for the multi-core systems illustrated in Fig. 1 and Fig. 2 in accordance with their equation numbers.

IV. DISTRIBUTED MEMORY PROGRAMMING

In this section, we briefly introduce the distributed memory programming concept. However, we first consider the sequential programming approach. In sequential programming, we use a single core for simulations. In our particular problem of solving linear algebraic equations, the sequential programming can be easily applied. We develop a sequential C program for solving linear algebraic equations distributively.

In this code snippet, we first define the problem of interest in the form of (1) in line 19. Then we defined the connection relationship between the nodes to utilize data produced by the neighboring node. Next, we directly use the algorithm defined in (2) for each node. In here, we used *rank* as the current processed node. All these processes are done by a single core and each node waits for others to finish their computations.

Now, we study the distributed memory programming. Essentially, a distributed memory system contains multiple core and local memory pairs connected by a network. Each core in this system processes its own work in accordance with its clock. However, when we need to send and receive data from one node to another to accomplish a complex task, we need to utilize a message passing protocol between the corresponding cores. One of the most used standard on distributed memory systems is MPI. We simulate the problems we defined in the previous section by using MPI on a distributed memory system.

We develop the program in Listing 2 by utilizing the MPI standard in C language to solve linear algebraic equations in a distributed manner. In the given program, we first initialize the MPI with the number of cores we have in the distributed

```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <math.h>
4
5  #define ERROR ... // tolerance
6  #define N ... // number of agents
7
8  // Prototypes
9  void problemDefinition();
10 int determineInNeighbor(int rank);
11 void algorithmMou(int rank, float x_neighbor[N]);
12 float norm(float soln[], float unique[], int
    ↪ num_rows);
13 void printVector(float vector[N], int size);
14
15 int main(int argc, char** argv) {
16 int rank;
17 float soln_norm;
18 float x_neighbor[N];
19 problemDefinition();
20
21 do{
22     for (rank = 0; rank < N ; rank++){
23         x_neighbor =
24             ↪ determineInNeighbor(rank);
25         algorithmMou(rank, x_neighbor);
26         soln_norm = norm(x_next[rank], x[rank], N);
27     }while (soln_norm > ERROR);
28 printVector(x_next[0], N*N);
29 return EXIT_SUCCESS;
30 }

```

Listing 1: A sequential program for solving LAEs.

TABLE I
COMPUTATION TIMES.

	LAE in (3)	LAE in (4)
Sequential Program	418 μ sec	36245 μ sec
Distributed-memory Program (MPI)	34169 μ sec	40223 μ sec

system in lines between 21 and 23. Then, *MPI_Bcast* in line 24 is used to broadcast initial data to all cores to be used in the system. Next, each core exchanges data between its neighboring cores defined in the network. It is important to note that, each core processes its received data simultaneously until the tolerance is achieved. Then *MPI_Gather* collects data from each core into the designated destination core in line 33.

V. RESULTS

In this section, we review the results for the LAEs defined Section III by using sequential and distributed-memory programs defined in Listing 1 and 2.

In Table I, we provided the computation times for solving LAEs in (3) and (4) under the network topologies illustrated in Fig. 1 and Fig. 2. The values given in the table are the average computation times of 150 simulations in the same computer.

As can be observed from the table, the computation time needed to solve the problem (3) with sequential program is explicitly less than we have with distributed-memory program. On the other hand, we achieved consensus for problem (4) almost at the same average time for both approaches.

These results mainly show that the distributed-memory system structure superior the sequential approach when the

```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <math.h>
4  #include <mpi.h>
5
6  #define ERROR ... // tolerance
7  #define N ... // number of agents
8
9  // Prototypes
10 void problemDefinition();
11 int determineInNeighbor(int rank);
12 int determineOutNeighbor(int rank);
13 void algorithmMou(int rank, float x_neighbor[N]);
14 float norm(float soln[], float unique[], int
    ↪ num_rows);
15 void printVector(float vector[N], int size);
16
17 int main(int argc, char** argv) {
18 int rank, size, tagno = 101, iter;
19 float soln_norm, result[N*N], x_neighbor[N];
20 problemDefinition();
21 MPI_Init(&argc, &argv);
22 MPI_Comm_rank(MPI_COMM_WORLD, &rank);
23 MPI_Comm_size(MPI_COMM_WORLD, &size);
24 MPI_Bcast(x_next, N * N, MPI_FLOAT, 0,
    ↪ MPI_COMM_WORLD);
25 do{
26     neighbor_in = determineInNeighbor(rank);
27     neighbor_out =
28         ↪ determineOutNeighbor(rank);
29     MPI_Send(&x_next[rank], N, MPI_FLOAT,
30         ↪ neighbor_out, tagno,
31         ↪ MPI_COMM_WORLD);
32     MPI_Recv(&x_neighbor, N, MPI_FLOAT,
33         ↪ neighbor_in, tagno, MPI_COMM_WORLD,
34         ↪ MPI_STATUS_IGNORE);
35     algorithmMou(rank,
36         ↪ x_neighbor);
37     soln_norm = norm(x_next[rank],
38         ↪ x_current[rank], N);
39 }while (soln_norm < ERROR);
40 MPI_Gather(&x_next[rank], N, MPI_FLOAT, &result[rank],
    ↪ N, MPI_FLOAT, 0, MPI_COMM_WORLD);
41 if (rank == 0) {
42     printVector(result, N*N);
43 }
44 MPI_Finalize();
45 return EXIT_SUCCESS;
46 }

```

Listing 2: A distributed-memory program for solving LAEs.

underlying topology gets complicated. The main reason for that is a node in distributed-memory system do not necessarily has to wait to process its data until the process of all the other nodes. On the contrary, each node must wait for the others in sequential approach since it uses a single core.

VI. CONCLUSION

In this paper, we explored distributed-memory programming with MPI for solving linear algebraic equations in a distributed manner. We also reviewed the sequential approach for the same problem of interest. We compared the simulation times with two different problems under different network topologies for both approaches. Although the network topology is getting complicated when we have a larger system, simulations on the distributed-memory systems may be preferable. For future work, it would be convenient to operate with more comprehensive multi-core computing systems such as super computers in TRUBA in order to simulate more complicated systems. Also, we would like to study other parallel programming techniques to assess the performance more accurately.

REFERENCES

- [1] Jin, H., Jaspersen, D., Mehrotra, P., Biswas, R., Huang, L. & Chapman, B. High performance computing using MPI and OpenMP on multi-core parallel systems. *Parallel Computing*. **37**, 562-575 (2011)
- [2] LeBlanc, T., Subhlok, J. & Gabriel, E. A high-level interpreted MPI library for parallel computing in volunteer environments. *2010 10th IEEE/ACM International Conference On Cluster, Cloud And Grid Computing*. pp. 673-678 (2010)
- [3] Imamura, T., Tsujita, Y., Koide, H. & Takemiya, H. An architecture of Stampi: MPI library on a cluster of parallel computers. *Recent Advances In Parallel Virtual Machine And Message Passing Interface: 7th European PVM/MPI Users' Group Meeting Balatonfured, Hungary, September 10-13, 2000 Proceedings 7*. pp. 200-207 (2000)
- [4] Bridges, M., Vachharajani, N., Zhang, Y., Jablin, T. & August, D. Revisiting the sequential programming model for the multicore era. *IEEE Micro*. **28**, 12-20 (2008)
- [5] Eijkhout, V. Teaching distributed memory programming from mental models. *Journal Of Parallel And Distributed Computing*. **118** pp. 120-127 (2018)
- [6] Hikosaka, O., Nakahara, H., Rand, M., Sakai, K., Lu, X., Nakamura, K., Miyachi, S. & Doya, K. Parallel neural networks for learning sequential procedures. *Trends In Neurosciences*. **22**, 464-471 (1999)
- [7] Mou, S., Liu, J. & Morse, A. A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions On Automatic Control*. **60**, 2863-2878 (2015)

Radyasyon Dedektörü Eklenmiş Otonom bir TurtleBot3 ile Radyasyon Haritalanması için bir ROS Paketi

A ROS Package for Radiation Mapping with an Autonomous TurtleBot3 Equipped with a Radiation Detector

İlayga AVAN
Elektrik-Elektronik Mühendisliği
İstanbul Sabahattin Zaim Üniversitesi
İstanbul, Türkiye
avan.ilyaga@std.izu.edu.tr

Oktay DOĞANGÜN
Bilgisayar Mühendisliği Bölümü
İstanbul Sabahattin Zaim Üniversitesi
İstanbul, Türkiye
oktay.dogangun@izu.edu.tr

Aydın Tarık ZENGİN
Elektrik Mühendisliği
İstanbul Teknik Üniversitesi
İstanbul, Türkiye
tarik.zengin@itu.edu.tr

Ferhat ÖZOK
Fizik Bölümü
Mimar Sinan Güzel Sanatlar Üniversitesi
İstanbul, Türkiye
ferhat.ozok@msgsu.edu.tr

M. Nizamettin ERDURAN
Nükleer Algılayıcılar ve Robotik Uygulama
ve Araştırma Merkezi
İstanbul Sabahattin Zaim Üniversitesi
İstanbul, Türkiye
nizamettin.erduran@izu.edu.tr

Taylan YETKİN
Biyomedikal Mühendisliği Bölümü
Yıldız Teknik Üniversitesi
İstanbul, Türkiye
taylan.yetkin@ytu.edu.tr

Özetçe— Bu çalışma, bir sensör olarak radyasyon sayacı donanımı eklenmiş otonom bir TurtleBot3 için açık kaynaklı ROS (Robot Operating System) tabanlı bir paket geliştirmeyi amaçlamaktadır. Bu paket, robotun kontrolünü, iletişimini ve çevrenin radyasyon haritalanmasını sağlamaktadır.

Hali hazırda bulunan bir radyasyon dedektörünün sensör olarak eklendiği robotun böcek algoritması sayesinde otonom hareket ederek radyasyon içeren belli bir alanı taraması ve radyasyon sayımı bilgisini yayınlaması sağlanmıştır.

Çalışmanın donanım tarafında, bir fotoçoğaltıcı tüpten yükseltilecek çıkan analog sinyaller, önce bir pencere ayırıcı devresi yardımıyla tek kanaldan dijital hale getirilmiştir. Ardından bir mikrodenetleyici sayesinde belli bir sürede alınan dijital sinyaller sayılarak TurtleBot3'te bulunan Raspberry pi üzerindeki USB portuna ulaştırılmıştır. Çalışmanın yazılım tarafında ise, mikrodenetleyiciden çıkan radyasyon ölçümü verisini seri porttan dinleyen ve bunu ROS topic olarak yayınlayan bir modül yazılmıştır.

Çalışmada yazılan ROS modüllerinin TurtleBot3 meta paketine uygun olması sayesinde sürü robot çalışmalarında kullanılması veya benzer platformlara uygulanması sağlanmıştır. Bu amaçla, radyasyon ölçümü verisinin ROS topic üzerinde yayını dinleyen ve SLAM tekniği ile bunu ısı haritası olarak yayınlayan bir modül eklenmiştir.

Daha sonraki çalışmalarda bu paketin birden çok dedektör olması durumu için geliştirilmeler yapılması ve böcek algoritmasından daha verimli alan tarama tekniklerinin eklenmesi planlanmaktadır.

Anahtar Kelimeler — Robot Operating System; TurtleBot3; Gama Dedektörü; Radyasyon Haritalama.

Abstract— This study aims to develop an open-source ROS-based package for an autonomous TurtleBot3 equipped with a

radiation counter hardware as a sensor, enabling control, communication of the robot, and radiation mapping of the environment.

By adding an existing radiation detector as a sensor, the robot, utilizing a bug algorithm, autonomously moves to scan an area containing radiation and publishes the radiation count information.

On the hardware side of the study, the analog signals generated by a photomultiplier tube are first converted to digital signal through a single-channel process using a window discriminator circuit. Then, the digital signals, counted over a specific period, are sent to the USB port on the Raspberry Pi mounted on the TurtleBot3 via a microcontroller. On the software side, a module was written to listen to the radiation measurement data from the microcontroller via the serial port and publish it as a ROS topic.

Since the ROS modules developed in this work are compatible with the TurtleBot3 meta-package, they can be used in swarm robot studies or applied to similar platforms. To achieve this, an additional module was added to listen to the radiation measurement data published on the ROS topic and broadcast it as a heatmap using SLAM techniques.

Further development of this package to handle multiple detectors and the addition of more efficient area scanning techniques beyond the bug algorithm might be studied in the future.

Keywords — Robot Operating System; TurtleBot3; Gamma Detectors, Radiation Mapping.

I. GİRİŞ

Bu çalışma, otonom bir TurtleBot3 robotuna bağlı bir radyasyon dedektörünü sensör olarak kullanarak radyasyon

haritalama yeteneğine sahip açık kaynaklı bir ROS (Robot Operating System – Robot İşletim Sistemi) paketi [2-7] geliştirmeyi amaçlamaktadır. Robot, çevresindeki radyasyonu ölçüp haritalayarak özellikle nükleer güvenlik, çevresel izleme ve radyasyon tespiti gibi alanlarda kullanılabilir bir sistem sunmaktadır.

Projenin donanım kısmında, gelen radyasyonun elektrik sinyaline dönüştürüldüğü bir foto-çoğaltıcı tarafından üretilen analog atımlar, *pencere ayırıştırıcı* (window discriminator) devresi kullanılarak dijital sinyallere dönüştürülmüştür. Bu dijital sinyaller, belirli bir zaman diliminde sayılarak bir mikrodenetleyici aracılığıyla TurtleBot3'ün üzerine monte edilmiş Raspberry Pi'ye USB portu üzerinden iletilmiştir. Yazılım kısmında ise, mikrodenetleyiciden gelen radyasyon ölçüm verilerini seri port üzerinden dinleyen ve bu verileri bir ROS konusu olarak yayınlayan bir modül geliştirilmiştir.

Bu sistemin diğer robotik uygulamalara kolayca entegre edilebilmesi için, geliştirilen ROS modülleri TurtleBot3 meta-paketleriyle uyumlu olacak şekilde tasarlanmıştır [1]. Buna ek olarak, radyasyon ölçüm verilerini bir ROS topic üzerinden alarak işleyen ve bu verileri SLAM (Simultaneous Localization and Mapping — Eşzamanlı Konumlandırma ve Haritalama) tekniklerini kullanarak ısı haritasına dönüştüren bir modül geliştirilmiştir. Bu modül sayesinde, robotun hareket ettiği alanlarda radyasyon dağılımı gerçek zamanlı olarak görselleştirilebilmektedir [13, 14, 17, 18].

Bu çalışma, ROS tabanlı modüllerin esnekliği ve TurtleBot3'ün görece düşük maliyetli ve sağlam platformu sayesinde, radyasyon algılama ve haritalama alanındaki mevcut çözümlere açık kaynaklı bir alternatif sunmaktadır. Gelecekte bu paketin birden fazla dedektörü destekleyecek şekilde genişletilmesi ve daha verimli alan tarama algoritmalarıyla zenginleştirilmesi planlanmaktadır. Bu yönüyle proje hem akademik hem de uygulamalı robotik çalışmalar için önemli bir katkı sunmaktadır.

A. ROS ve Robotik Sistemler

ROS, robotik uygulamalar için tasarlanmış açık kaynaklı bir yazılım çerçevesidir ve robotların farklı sensör ve aktüatörlerle entegre bir şekilde çalışmasını sağlar [2,3,5,7]. ROS, düğüm (node) yapısına dayanarak farklı görevlerin bağımsız modüller üzerinden gerçekleştirilmesine olanak tanır. Bu çalışma kapsamında, ROS'un sunduğu iletişim altyapısı ve modül yapı, radyasyon verilerinin toplanması, işlenmesi ve görselleştirilmesi için kullanılmıştır.

ROS, robotik yazılım geliştirmeyi kolaylaştırmayı hedefleyen açık kaynaklı bir platformdur. Robotik sistemlerin tasarımı, geliştirilmesi ve simülasyonu için gerekli olan araçları sunar. Kullanım alanı oldukça geniş olan ROS, endüstriyel robotlar, servis robotları, otonom araçlar ve akademik araştırma projeleri gibi pek çok alanda uygulanabilir.

Bu platform, endüstriyel robotların kontrolü, izlenmesi ve programlanması, servis robotlarının temizlik, teslimat ve hasta bakımı gibi görevleri üstlenmesi, otonom araçların navigasyon, algılama ve karar verme süreçlerinde önemli bir rol oynar. Aynı zamanda, robotik derslerinde, akademik çalışmalarda ve bireysel projelerde de sıklıkla tercih edilmektedir.

Modül yapısı sayesinde farklı bileşenlerin kolayca entegre edilmesini sağlayan ROS, dağıtılmış mimarisiyle

cihazlar arasında mümkün olduğunca sorunsuz iletişim kurulmasını mümkün kılar. Önceden hazırlanmış modüllerin yeniden kullanılabilir olması, geniş bir kullanıcı ve geliştirici topluluğuna sahip olması, güçlü simülasyon ve görselleştirme araçları sunması ve hata ayıklama süreçlerini basitleştirmesi, ROS'u robotik alanında öne çıkan vazgeçilmez bir açık kaynak platform haline getirmektedir.

Bu çalışmada kullanılan TurtleBot3 robot platformunun açık kaynaklı olarak yazılmış paketlerinin kararlı olarak çalıştığı bir sürüm olması nedeniyle ROS'un en yaygın sürümlerinden olan noetic tercih edilmiştir.

B. Radyasyon Dedektörü ve Sinyal İşleme

Radyasyon dedektörleri, nükleer ve parçacık fizik deneylerinde, çevresel radyasyon izlemelerinde ve güvenlik uygulamalarında yaygın olarak kullanılmaktadır [21, 22]. Bu çalışmada kullanılan dedektör, yarı iletken foto-çoğaltıcı ile çalışmaktadır. Foto-çoğaltıcı, radyasyonun neden olduğu ışık sinyallerini elektrik sinyallerine dönüştürür ve bu sinyallerin yoğunluğunu ölçer.

Analog sinyallerin işlenmesi için bir *pencere ayırıştırıcı* devresi (window discriminator) kullanılmıştır. Bu devre, sinyalleri belirli bir alt ve üst eşik voltajı arasında filtreleyerek gürültüyü azaltır ve yalnızca istenen aralıktaki sinyalleri dijital sinyallere dönüştürür.

Dijital sinyaller, bir mikrodenetleyici (Arduino) tarafından sayılarak TurtleBot3 üzerinde yer alan Raspberry Pi'ye iletilir. Mikrodenetleyicinin sunduğu *kesinti* (interrupt) mekanizması, her bir sinyalin hızlı ve doğru bir şekilde sayılmasını sağlar. Bu sayılar, ROS ortamında işlenerek radyasyon yoğunluğuna ilişkin bilgiler elde edilir.

C. SLAM ve Isı Haritalama

SLAM, bir robotun bulunduğu ortamın haritasını çıkarırken aynı anda kendi konumunu belirlemesine olanak tanır [9-12, 15]. Bu çalışmada, TurtleBot3'ün lazer mesafe ölçer ve odometre verileri, çevrenin haritasını çıkarmak için kullanılmıştır. Radyasyon dedektöründen alınan veriler, SLAM haritası üzerine işlenerek bir ısı haritası oluşturulmuştur.

Isı haritaları, çevresel değişkenlerin mekânsal dağılımını görselleştirmek için kullanılan etkili bir yöntemdir. Bu çalışmada radyasyon yoğunluğunu temsil etmek için renk skalası kullanılan bir ısı haritası oluşturulmuştur. Robotun bulunduğu konuma bağlı olarak, dedektör verileri belirli hücrelere aktarılmış ve gerçek zamanlı olarak güncellenmiştir. Bu yöntem, radyasyon kaynaklarının tespiti ve yoğunluklarının görselleştirilmesi açısından önemli bir avantaj sağlamaktadır. SLAM (Simultaneous Localization and Mapping), robotların bilinmeyen bir ortamda hem kendi konumlarını belirlemesine hem de bu ortamın haritasını çıkarmasına olanak tanıyan bir tekniktir. Bu teknoloji, robotik sistemlerde otonomi sağlamada kritik bir rol oynar ve GPS'in yetersiz kaldığı iç mekân gibi alanlarda geniş uygulama alanına sahiptir. SLAM, lazer tarayıcılar, kamera sistemleri, IMU (Inertial Measurement Unit) ve odometre gibi sensörlerden alınan verileri birleştirerek çevrenin dinamik bir haritasını oluşturur ve robotun çevresine adaptasyonunu mümkün kılar.

Isı haritalama ise, mekânsal bir bölgede belirli bir değişkenin (örneğin sıcaklık, nem veya radyasyon) yoğunluk dağılımını görselleştirmek için kullanılan bir tekniktir. Isı haritaları, genellikle farklı yoğunlukları renk skalasıyla

temsil ederek karmaşık verileri kolayca yorumlanabilir hale getirir. Radyasyon yoğunluğu gibi risk taşıyan değişkenlerin tespitinde, bu yöntem hem güvenlik önlemlerinin alınması hem de ortam analizi için değerli bilgiler sağlar.

SLAM ve ısı haritalamanın birleşimi, mobil robotların çevresel veri toplama ve analizinde güçlü bir araçtır. Örneğin, bir robot, SLAM sayesinde ortamın detaylı bir haritasını oluştururken, aynı anda belirli sensörlerden alınan verileri bu harita üzerine işleyebilir. Bu, özellikle tehlikeli bölgelerde (nükleer tesisler, madenler veya doğal afet alanları) görev yapan robotlar için kritik öneme sahiptir. Ayrıca, bu entegrasyon, çevresel değişikliklerin gerçek zamanlı olarak izlenmesi ve tespit edilmesi için büyük bir avantaj sunar.

D. TurtleBot3 Paket Yapısı ve ROS Topic İlişkileri

TurtleBot3, açık kaynaklı bir robot platformudur ve özellikle robotik uygulamaları öğrenmek ve geliştirmek isteyenler için popüler bir araçtır. Robotun donanımı, ROS (Robot Operating System) ile uyumlu olup, eğitim, araştırma ve prototip geliştirme alanlarında yaygın olarak kullanılır. TurtleBot3, farklı sensörler, motorlar ve bir bilgisayar sistemi ile donatılmıştır ve kullanıcılara robotik algoritmalar geliştirme, simülasyonlar yapma ve robotun hareketlerini kontrol etme olanağı sunar. Ayrıca, robot, kullanıcıların 3D haritalama, otonom navigasyon, nesne tanıma ve daha birçok robotik uygulama geliştirmelerine olanak tanır. TurtleBot3, modüler yapısı sayesinde farklı konfigürasyonlarda kullanılabilir ve robotik araştırmalar için güçlü bir başlangıç platformudur.

TurtleBot3, ROS tabanlı bir mobil robot platformudur ve modüler yapısı sayesinde geniş bir uygulama yelpazesine sahiptir. Bu platformda yer alan yazılım paketi yapısı, farklı görevlerin bağımsız olarak çalışmasını ve ROS ortamında kolayca entegrasyonunu sağlar. Geliştirilen `gamma_detection` paketi, TurtleBot3'ün mevcut yazılım altyapısına entegre edilerek radyasyon haritalama işlemini gerçekleştirmek için tasarlanmıştır.

TurtleBot3'ün ROS tabanlı yazılım yapısı aşağıdaki temel bileşenlerden oluşur:

- `turtlebot3_bringup`: Robotun temel bileşenlerini (sensörler, motorlar, vb.) çalıştırmak için kullanılan bir başlatma (launch) dosyasını içerir.
- `turtlebot3_description`: Robotun URDF (Unified Robot Description Format) dosyasını ve ilgili 3D modellerini içerir.
- `turtlebot3_slam`: Robotun SLAM algoritmalarını çalıştırarak harita oluşturmaya olanak tanır.
- `turtlebot3_navigation`: Robotun planlama ve otonom hareket yeteneklerini sağlar.

Kullanıcılar tarafından eklenen paketler, TurtleBot3'ün işlevselliğini artırmak için kullanılır. Bu çalışmada geliştirilen `gamma_detection` paketi, radyasyon algılama ve haritalama için yazılmıştır.

TABLE I. TURTLEBOT3 META PAKETİNİN DÜĞÜMLERİ.

Düğüm	Yayın	Abonelik	Görev
-------	-------	----------	-------

<code>turtlebot3_robot_state_publisher</code>	/tf	-	Robotun eklem durumlarını ve pozisyonunu yayımlar.
<code>turtlebot3_core</code>	/odom, /joint_states, /battery_status, /sensor_status	/cmd_vel	Motor ve sensör verilerini işler; odometri ve sensör durumlarını yayımlar.
<code>turtlebot3_slam_gmapping</code>	/map, /tf, /scan, /particlecloud	/odom, /tf, /scan	SLAM algoritmasını çalıştırarak bir harita oluşturur.
<code>turtlebot3_navigation</code>	/move_base/goal, /move_base/status, /map	/odom, /map, /scan, /move_base_simple/goal	Otonom navigasyon için rota planlaması ve hareket kontrolü sağlar.
<code>turtlebot3_teleop</code>	/cmd_vel	-	Klavye veya joystick girişiyle robotun hareketini kontrol etmek için hız komutları gönderir.
<code>turtlebot3_rviz</code>	-	/map, /odom, /scan, /tf	RViz üzerinden robotun durumunu, haritasını ve sensör verilerini görselleştirir.
<code>hlds_laser_publisher</code>	/scan	-	Lidar verilerini yayımlar.
<code>move_base</code>	/move_base/goal, /cmd_vel	/map, /odom, /scan, /tf	Hedef pozisyonlara gitmek için yol planlaması ve hız kontrolü yapar.

II. MALZEME VE YÖNTEM

Çalışmanın gerçekleştirileceği cihazlardaki sistem, kurulan kütüphane ortamı ve haberleşme düzeni için gerekli ayrıntılar, ardından çalışmanın konusu olan ROS paketinin ayrıntıları ve ayrıca robota eklenen donanımın ilgili açıklamaları bu bölümde verilecektir.

Çalışmada TurtleBot3, böcek algoritması kullanılarak belli bir alan içerisinde otonom hareket etmektedir. Robot önce rastgele bir hedef belirleyip böcek algoritması ile o hedefe doğru ilerleyecek şekilde çalıştırılmıştır. Eğer hedefe ulaşırsa yeniden rastgele bir konum belirleyerek hareketine devam eder. Her tekrarda alanın rastgele bir kısmını taramış olur.

Robot, alanı gezerken gama dedektörü ile alınan analog sinyal bir ayrıştırıcı devresi üzerinden dijital sinyale dönüştürülür. Ardından bir mikrodenetleyici yardımıyla her saniye gelen dijital atmalar sayılır ve seri port üzerinden TurtleBot3 üzerindeki Raspberry Pi 3'e yollanır. Cihazın üzerinde geliştirilmiş olan `turtlebot3_gamma_detection`

paketi, önce saniyelik radyasyon oranını `count_rate` olarak yayınlar. Bu yayından hareketle paketin içindeki ikinci bir düğüm tarafından bir ısı haritası oluşturup `heat_map` olarak yayınlanır (Figür 1).

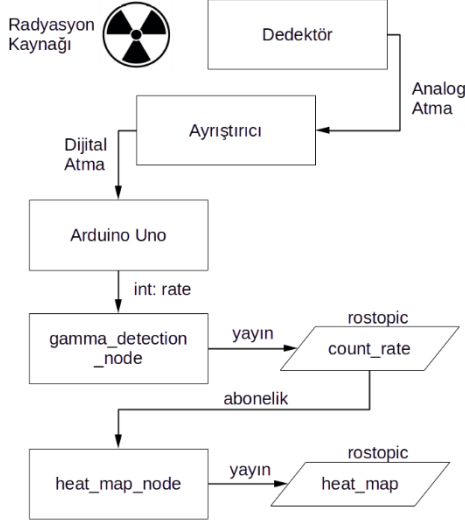


Fig. 1. Sistemin yapısı.

A. Sistemin kurulması

TurtleBot3 içinde bulunan ARM tabanlı işlemciye sahip Raspberry Pi 3 üzerine Ubuntu 20.04 kurulmuştur. Ardından robotların veya IoT (Internet of Things) cihazların haberleşmesini sağlamak amacıyla yazılmış açık kaynaklı ROS kütüphane ortamının Noetic sürümü, en kararlı olarak Ubuntu'nun söz konusu sürümünde çalışması nedeniyle, standart adımlar takip edilerek kurulmuştur.

Standart kurulum adımlarının bir parçası olarak TurtleBot3 ile haberleşmeyi ve uzaktan kontrolü sağlamak için, bundan sonra adına *Ana Bilgisayar* denilecek olan bir bilgisayara, yine Ubuntu 20.04 ve ROS Noetic kurulumu gerçekleştirilmiştir. Kurulumların sonunda ise aynı yerel ağda çalıştırılan Ana Bilgisayar ile TurtleBot3'lerin haberleşmesi için gerekli IP adresleri sabitleştirilerek her iki sisteme de tanımlanmıştır.

ROS, kullanıcıların çeşitli paketler yazması ile özel amaçlara hizmet edecek şekilde tasarlanmıştır. Kullanıcılar tarafından eklenen paketler, TurtleBot3'ün veya IoT cihazların işlevselliğini artırmak için kullanılır. Bu çalışmada geliştirilen `turtlebot3_gamma_detection` adlı paket, radyasyon algılama ve haritalama için TurtleBot3 robotuna ek paket olarak `turtlebot3` meta paketinin altında olacak şekilde yazılmıştır.

Söz konusu paketin iki düğümü (node) bulunur: radyasyon verisinin toplanmasını sağlayan `gamma_detection_node` düğümü ile haritalanacak veriyi sağlayan `heat_map_node` düğümü. Paketin başlatma (launch) dosyası, bu düğümlerin ikisinin de başlatılmasını sağlayacaktır.

B. Anlık dedektör verisinin toplanması

Bu bölümde, `gamma_detection_node` düğümünün işleyişi ayrıntılı olarak açıklanmıştır. Bu düğüm, TurtleBot3 robotuna entegre edilmiş radyasyon dedektöründen işlenerek gelen sayısal verileri bir mikrodenetleyici üzerinden alır ve ROS ortamında bir topic olarak yayınlar.

Sistem, ROS tabanlı modüler bir yapı ile tasarlanmıştır ve donanım entegresi ile düğüm yapısı olarak iki ana bileşen içerir. Radyasyon dedektöründen alınan analog sinyaller, bir pencere ayırmacı (window discriminator) devresi kullanılarak dijital sinyallere dönüştürülür. Mikrodenetleyici (Arduino), bu dijital sinyalleri sayar ve belirli bir zaman dilimi için biriken radyasyon ölçümünü seri port üzerinden Raspberry Pi'ye iletir. ARM tabanlı işlemciye sahip ve içine Ubuntu 20.04 kurulmuş olan Raspberry Pi, TurtleBot3 platformuna entegre edilmiştir ve ROS düğümlerini çalıştırır.

Donanım olarak entegre edilen radyasyon dedektörü, bir foto-çoğaltıcının önüne yerleştirilmiş inorganik bir kristalden oluşmaktadır (Figür 2). Çalışmamızda silikon tabanlı foto-çoğaltıcı olarak SensL markasının J serisi 4mm x 4mm boyutundaki tekli SiPM (Silicon Photo-Multiplier) ürünü kullanılmıştır. Foto-çoğaltıcının aktif alanına iliştirilmiş olan 5mm×5mm×2inç boyutlarındaki LYSO (Lutetium-yttrium oxyorthosilicate) kristali, gelen gama ışınlarının etkileştiği ve görünür dalga boyundaki ışığa dönüştüğü kısmı oluşturur.



Fig. 2. Yarı iletken foto-çoğaltıcı ile LYSO kristalinin birleştirilip, ön yükselteç devresi üzerinden çıkış alacak ve besleme girişleri bulunacak şekilde ışık geçirmez bir kutuya yerleştirilmiş dedektör.

Çalışmada kullanılan söz konusu spesifik dedektörün seçilmesinin nedeni, hali hazırda ucuza mal edilmiş bir dedektör olmasının yanında vakum tüp içinde fotoelektrik etkiye dayalı olan foto-çoğaltıcı tüp yerine yarı iletken teknolojisine dayanan SiPM olması nedeniyle daha hafif ve az yer kaplamasıdır. Öte yandan küçük aktif alanı, çalışmada gelen radyasyonun enerji ölçümü yapılmadığı ve sadece gamaların sayımı önemli olduğu için bir dezavantaj yaratmamaktadır.

Foto-çoğaltıcı bu görünür dalga boyundaki ışığa duyarlı fotodiyotlardan meydana geldiği için çıkış olarak onlarca mV mertebesinde gerilim genliğine sahip analog bir elektrik sinyali üretir. Sinyal çıkışına bağlanmış bir ön yükseltici devresinin gönderdiği volt mertebesine yükselttilip şekillendirilmiş analog sinyaller, karakteristik olarak hızlı bir çıkışa ve görece daha yavaş bir inişe sahip olan bir atma biçimindedir. Atmalar ön yükselticiden geçirmiş durumda birkaç yüz nanosaniye kadarlık sürelerde oluşmaktadır. Bu süre, sinyal ayırmacı bir devreyi tetikleyebilecek uzunluktadır.

Analog olan atma sinyallerin genliği bu tür dedektörlerde gelen gamanın enerjisiyle doğru orantılıdır. Ancak bu çalışmada gamaların enerjisi ölçülmeyecek olup yalnızca sayılması yeterli olduğu için dijital ayırmacı devrede ortamda bir radyasyon kaynağının olmadığı durumdaki ardalan radyasyondan oluşan istenmeyen atmalardan ayrıştırılması üzerine odaklanılmıştır. Bu istenmeyen atmalar, ardalan gürültüsü olarak adlandırılır ve elektronik gürültüden daha yüksek genliklerde olabilir.

Ardalan gürültüden ve elektronik gürültüden ayrıştırmak için yapılan ayırmacı devresi, temel olarak tek kanal dijital

çıkışı olan, atmanın genliğinin belli bir değerden yukarıda olması durumunda 5V genlikte olan TTL HIGH (mantıksal 1 değerinde *transistor-to-transistor logic* standardı) dijital bir atma sağlar, onun dışındaki durumlarda ise yaklaşık 0V genlikte olan TTL LOW (mantıksal 0 değerinde *transistor-to-transistor logic* standardı) durumunda kalır.

Ayrımcı devre, gelen analog atmanın genliği aradalan gürültüye göre ayarlanan iki değerin arasında olduğunda tetiklenir ve tetiklendikten sonra belli bir süre boyunca dijital bir kare dalga oluşturur. Oluşan kare dalganın uzunluğu, analog sinyalin uzunluğuyla doğru orantılı olup birkaç 100ns sürede üretilmektedir. Bu süre, Arduino gibi bir mikrodenetleyicinin dijital kesinti pinlerinden okunabilecek bir süredir ve bu sayede mikrodenetleyicideki kod istenilen işlemleri gerçekleştirecek şekilde tetiklenir.

Mikrodenetleyicideki kesinti pini, dijital TTL HIGH atması gelmesi durumunda tetiklenip C dilinde yazılıp derlenmiş kodda söz konusu pine atanmış olan fonksiyon çalışır. Bu fonksiyon, her kesinti olduğunda global bir değişkeni +1 artırır. Mikrodenetleyicide sürekli olarak çalışan loop() döngüsü ise tutulan değişkeni her saniye seri porta yazıp değerini sıfırlar.

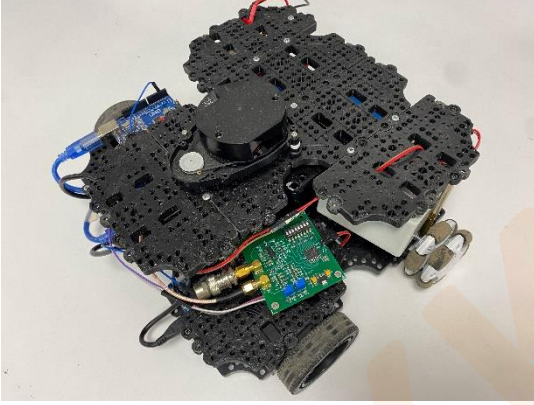


Fig. 3. Radyasyon dedektörü eklenmiş TurtleBot3.

C. Radyasyon sıklığının konuma göre yayınlanması

Çalışmada geliştirilen `turtlebot3_gamma_detection` paketi, SLAM haritası üzerine radyasyon verilerini işlemek için aşağıdaki adımları izler:

- SLAM haritasından alınan robot konumu (`/odom` topic'i üzerinden).
- Radyasyon verilerinin (`/count_rate`) robotun bulunduğu harita hücresine işlenmesi.
- Haritanın güncellenerek `/heat_map` topic'i üzerinden yayınlanması.

Bu yapı sayesinde TurtleBot3, SLAM ile elde edilen harita üzerinde radyasyon yoğunluklarını gerçek zamanlı olarak görselleştirebilir. Kullanıcı, RViz üzerinde `/heat_map` topic'ini bir katman olarak ekleyerek radyasyon dağılımını renkli bir ısı haritası olarak gözlemleyebilir.

Bu bölümde, `gamma_detection_node` düğümünün işleyişi detaylandırılmıştır. Bu düğüm, TurtleBot3 robotuna entegre edilmiş radyasyon dedektöründen gelen verileri alır ve ROS ortamında bir topic olarak yayınlar.

Radyasyon dedektöründen alınan analog sinyaller, bir pencere ayrımcısı (*window discriminator*) devresi kullanılarak dijital sinyallere dönüştürülür. Mikrodenetleyici (Arduino), bu dijital sinyalleri sayar ve belirli bir zaman dilimi için biriken radyasyon ölçümünü seri port üzerinden Raspberry Pi'ye iletir. Bu noktada `gamma_detection_node`, seri port üzerinden gelen verileri dinler ve her yarım saniyede bir alınan sayım oranını (`count_rate`) `std_msgs/Int32` türünde bir ROS topic olarak yayınlar. Düğümün işleyişi şu şekilde özetlenebilir:

- Seri porttan gelen veri okunur ve çözümlenir.
- Alınan radyasyon ölçüm değeri bir ROS topic olan `/count_rate` üzerinden yayınlanır.
- Diğer düğümlerin bu verilere abone olmasına olanak sağlanır.

Burada `heat_map_node` paketin içinde bu yayına abone olarak SLAM haritası oluşturur ve `heat_map` olarak yayınlar. Bu yayın, `rviz` gibi görsel ara yüz paketleri ile görselleştirilebilir (Figür 4)

TABLE II. GELİŞTİRİLEN TURTLEBOT3 PAKETİNİN DÜĞÜMLERİ

Düğüm	Yayın	Abonelik	Görev
<code>gamma_detection_node</code>	<code>/count_rate</code>	-	Arduino'dan alınan verileri yayımlar.
<code>heat_map_node</code>	<code>/heat_map</code>	<code>/count_rate</code> <code>/odom</code>	Radyasyon verilerini haritalandırır.

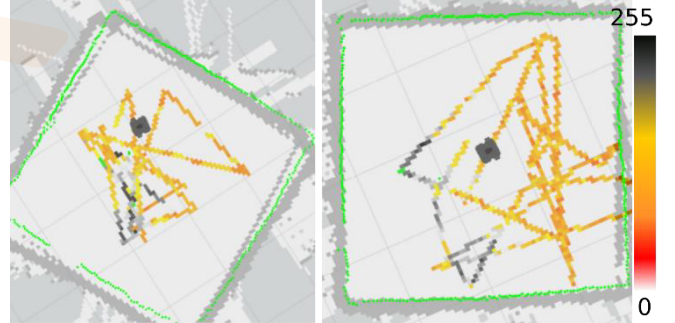


Fig. 4. Oluşan radyasyon haritası.

III. SONUÇ VE TARTIŞMA

Bu çalışmada, TurtleBot3 platformu üzerine geliştirilen `turtlebot3_gamma_detection` paketi kullanılarak, çevredeki radyasyon seviyelerinin dinamik olarak haritalandırılması gerçekleştirilmiştir. Radyasyon ölçümleri ve robotun pozisyon verileri başarıyla entegre edilerek, gerçek zamanlı bir "ısı haritası" elde edilmiştir.

Geliştirilen sistem, radyasyon seviyelerini saniyelik olarak ölçmüş ve robotun anlık konum bilgisi ile birleştirilerek 2D bir radyasyon haritası oluşturmuştur. Sistem, TurtleBot3'ün SLAM altyapısı ile tamamen uyumlu çalışmış ve ROS'un sağladığı araçlarla veriler başarılı bir şekilde işlenip görselleştirilmiştir.

Dedektörün çözünürlüğü artırılarak daha hassas ölçümler yapılabilir. Daha büyük alanlarda haritalama için robotun hareket planlama algoritmaları geliştirilebilir.

Bu sistem, özellikle nükleer santraller, maden sahaları ve nükleer malzeme depolarında radyasyon tespiti ve haritalandırılması için kullanılabilir. Ayrıca, akademik araştırmalar ve eğitim amaçlı uygulamalarda da faydalı bir platform sunmaktadır

Sonuç olarak, geliştirilen sistem, düşük maliyetli ve taşınabilir bir çözüm sunmakta olup, radyasyon haritalandırma çalışmalarında başarılı bir şekilde kullanılabilir. Gelecekte yapılacak iyileştirmelerle birlikte, sistemin endüstriyel ve araştırma amaçlı uygulamalarda daha geniş bir kullanım alanına sahip olacağı öngörülmektedir.

TEŞEKKÜR

Bu çalışma, İstanbul Sabahattin Zaim Üniversitesi BAP-2022-30 numaralı proje kapsamında Nükleer Algılayıcılar ve Robotik Uygulama ve Araştırma Merkezi (İZÜNAR) ve Mimar Sinan Güzel Sanatlar Üniversitesi Fizik bölümü olanaklarıyla gerçekleştirilmiştir.

KAYNAKLAR

- [1] Amsters, R., & Slaets, P. (2020). Turtlebot 3 as a robotics education platform. In *Robotics in Education: Current Research and Innovations 10* (pp. 170-181). Springer International Publishing.
- [2] Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., and Ng, A. Y. (2009). ROS: an open-source Robot Operating System. In *ICRA workshop on open source software* (Vol. 3, No. 3.2, p. 5).
- [3] Quigley, M., Gerkey, B., & Smart, W. D. (2015). *Programming Robots with ROS: a practical introduction to the Robot Operating System*. "O'Reilly Media, Inc."
- [4] Joseph, L. (2015). *learning Robotics using python*. Packt Publishing Ltd.
- [5] Furrer, F., Burri, M., Achtelik, M., and Siegwart, R. (2016). *Robot operating system (ros): The complete reference (volume 1)*. Cham: Springer International Publishing, 595-625.
- [6] Fairchild, C., & Harman, T. L. (2016). *ROS robotics by example*. Packt Publishing Ltd.
- [7] Koubaa, A. (Ed.). (2017). *Robot Operating System (ROS) (Vol. 1, pp. 112-156)*. Cham, Switzerland: Springer.
- [8] Joseph, L. (2018). *Robot operating system (ros) for absolute beginners (Vol. 2018)*. Berlin/Heidelberg, Germany: Springer.
- [9] Younes, G., Asmar, D., Shamma, E., & Zelek, J. (2017). Keyframe-based monocular SLAM: design, survey, and future directions. *Robotics and Autonomous Systems*, 98, 67-88.
- [10] Khairuddin, A. R., Talib, M. S., & Haron, H. (2015, November). Review on simultaneous localization and mapping (SLAM). In *2015 IEEE international conference on control system, computing and engineering (ICCSCE)* (pp. 85-90). IEEE.
- [11] Li, J., Cheng, L., Wu, H., Xiong, L., & Wang, D. (2012, June). An overview of the simultaneous localization and mapping on mobile robot. In *2012 Proceedings of International Conference on Modelling, Identification and Control* (pp. 358-364). IEEE.
- [12] Chiem, V. D. K. (2024). Implementing SLAM and Autonomous Navigation on a TurtleBot3.
- [13] Gunarso, M. A., & Tamba, T. A. (2024, October). A Study on the Navigation and Control Methods for Turtlebot3 Robot Using ROS. In *2024 14th International Conference on System Engineering and Technology (ICSET)* (pp. 192-197). IEEE.
- [14] Placed, J. A., Strader, J., Carrillo, H., Atanasov, N., Indelman, V., Carlone, L., & Castellanos, J. A. (2023). A survey on active simultaneous localization and mapping: State of the art and new frontiers. *IEEE Transactions on Robotics*, 39(3), 1686-1705.
- [15] Khairuddin, A. R., Talib, M. S., & Haron, H. (2015, November). Review on simultaneous localization and mapping (SLAM). In *2015 IEEE international conference on control system, computing and engineering (ICCSCE)* (pp. 85-90). IEEE.
- [16] Panigrahi, P. K., & Bisoy, S. K. (2022). Localization strategies for autonomous mobile robots: A review. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 6019-6039.
- [17] Stachniss, C., & Burgard, W. (2003, August). Exploring unknown environments with mobile robots using coverage maps. In *IJCAI (Vol. 2003, pp. 1127-1134)*.
- [18] Amigoni, F., & Caglioti, V. (2010). An information-based exploration strategy for environment mapping with mobile robots. *Robotics and Autonomous Systems*, 58(5), 684-699.
- [19] Escobar-Naranjo, J., Caiza, G., Ayala, P., Jordan, E., Garcia, C. A., & Garcia, M. V. (2023). Autonomous navigation of robots: optimization with DQN. *Applied Sciences*, 13(12), 7202.
- [20] Xu, Q., Zhao, J., Zhang, C., & He, F. (2015, August). Design and implementation of an ROS based autonomous navigation system. In *2015 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 2220-2225). IEEE.
- [21] Zakaria, A. H., Mustafah, Y. M., Abdullah, J., Khair, N., & Abdullah, T. (2017). Development of autonomous radiation mapping robot. *Procedia Computer Science*, 105, 81-86.
- [22] Baca, T., Jilek, M., Manek, P., Stibinger, P., Linhart, V., Jakubek, J., & Saska, M. (2019, November). Timepix radiation detector for autonomous radiation localization and mapping by micro unmanned vehicles. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1129-1136). IEEE.
- [23] Karalekas, G., Vologiannidis, S., & Kalomiros, J. (2020). Europa: A case study for teaching sensors, data acquisition and robotics via a ROS-based educational robot. *Sensors*, 20(9), 2469.

