

T.C.
İSTANBUL SABAHATTİN ZAİM ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR BİLİMLERİ VE MÜHENDİSLİĞİ (%30 İngilizce)
BİLİM DALI

GENİŞLETİLMİŞ EVRİŞİMLİ SINIR AĞLARI
KULLANARAK GERÇEK ZAMANLI YÜZ TANIMA

DOKTORA TEZİ

Mehmet Ali KUTLUGÜN

İstanbul
Nisan - 2023

T.C.
İSTANBUL SABAHATTİN ZAİM ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR BİLİMLERİ VE MÜHENDİSLİĞİ (%30 İngilizce)
BİLİM DALI

GENİŞLETİLMİŞ EVRİŞİMLİ SİNİR AĞLARI KULLANARAK
GERÇEK ZAMANLI YÜZ TANIMA

DOKTORA TEZİ

Mehmet Ali KUTLUGÜN

Tez Danışmanı
Dr. Yahya ŞİRİN

İstanbul
Nisan - 2023

TEZ ONAYI

Lisansüstü Eğitim Enstitüsü Müdürlüğüne,

Bu çalışma, jürimiz tarafından Bilgisayar Bilimleri ve Mühendisliği Anabilim Dalı, Bilgisayar Bilimleri ve Mühendisliği Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman: Dr. Yahya ŞİRİN

Üye Dr. Öğr. Üyesi Mehtap YALÇINKAYA

Üye Dr. Öğr. Üyesi Abdullah SÖNMEZ

Üye Dr. Öğr. Üyesi, Şengül BAYRAK HAYTA

Üye Dr. Öğr. Üyesi Sümeyra BEDİR

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

Doç. Dr. Erhan İÇENER
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİMİ

Doktora tezi olarak hazırladığım “**Geniřletilmiř Evriřimli Sinir Ađları Kullanarak Gerçek Zamanlı Yüz Tanıma**” adlı alıřmanın neri ařamasından sonulandıđı ařamaya kadar geen srete bilimsel etiđe ve akademik kurallara zenle uyduđumu, tez iindeki tm bilgileri bilimsel ahlak ve gelenek erevesinde elde ettiđimi, tez yazım kurallarına uygun olarak hazırladıđımı, bu alıřmamda dođrudan veya dolaylı olarak yaptıđım her alıntıya kaynak gsterdiđimi ve yararlandıđım eserlerin kaynakada gsterilenlerden oluřtuđunu beyan ederim.

Mehmet Ali KUTLUGN

ÖN SÖZ

Araştırmanın yapılması sırasında yazdığımız ve kabul edilen Federated Conference on Computer Science and Information Systems (FedCSIS 2019) konferans bildirisinin sunumu için verdikleri destekten ötürü Türk Hava Yolları Teknoloji ve Bilişim A.Ş. Ar-Ge Merkezi grubuna teşekkür ederim. Kurum yetkililerine şükranlarımı sunarım.

Tez çalışmamın gerçekleşmesi aşamasında beni yönlendiren ve tecrübelerini paylaşan tez danışmanı hocam Dr. Yahya ŞİRİN'e teşekkür ederim. Tez izleme komitesinde bulunan kıymetli hocalarım Dr. Mehtap YALÇINKAYA ve Dr. Abdullah SÖNMEZ'e şükranlarımı sunarım.

Ayrıca kaynaklarından faydalandığım tüm akademik eserlerin yazarlarına katkılarından ötürü teşekkür ederim.

Mehmet Ali KUTLUGÜN
İstanbul - 2023

ÖZET
GENİŞLETİLMİŞ EVRİŞİMLİ SİNİR AĞLARI KULLANARAK
GERÇEK ZAMANLI YÜZ TANIMA

Mehmet Ali KUTLUGÜN

Doktora, Bilgisayar Bilimleri ve Mühendisliği (%30 İngilizce)

Tez Danışmanı: Dr. Yahya ŞİRİN

Nisan, 2023 - 98 +XV Sayfa

Bu çalışmanın konusu, gerçek zamanlı görüntülerde derin öğrenme temelli yüz tanıma sistemleri üzerinedir. Araştırma, yüz tanıma sistemleri içerisinde öznitelik elde etmek için kullanılan evrişimli sinir ağları kapsamında gerçekleştirilmiştir. Gerçek zamanlı görüntüler, “yüz bulma”, “öznitelik çıkarma”, “vektörel dönüşüm”, “sınıflandırma ve başarımlı ölçme” olmak üzere dört aşamada incelenmiştir. Yüz tanıma sistemlerinde, ışık yönü, yansıma, yüz ifadesindeki duygusal ve fiziksel değişimler tanımayı zorlaştıran başlıca unsurlar olarak bilinmektedir. Akan görüntü çerçevelerinde karşılaştırılan her görüntüde meydana gelebilen bu unsurlar hatalı tanımlamalara yol açmaktadır. Araştırmada, arka plan veya sahne değişimleri tespit yöntemi önerilerek, evrişimli sinir ağlarından elde edilen özniteliklerin karşılaştırması birlikte kullanılmıştır. Bu model, sahnenin değişmediği blokları tespit eder ve sınıflandırıcı aşamasında kullanılan karşılaştırma eşik değerini yeni bir değerle daha kesin olarak belirlemeye çalışır. Değişmeyen sahne blokları arasında hassasiyet oranının artırılması, veri tabanındaki örnekler arasında daha az karşılaştırma yapılmasına olanak tanır. Deneysel çalışmada önerilen model, orijinal derin metrik öğrenme modeline kıyasla %99,25 doğruluk ve %99,28 F-1 değerlerine ulaşmıştır. Deneysel sonuçlar, değişmeyen sahnelerde aynı kişinin yüz görüntülerinde farklılıklar olsa bile, karşılaştırılan örneklem alanı daraltıldığı için yanlış tanımanın çok daha aza indirilebileceğini göstermektedir.

Anahtar Kelimeler: Derin metrik öğrenme, gerçek zamanlı yüz tanıma, görüntü işleme, öznitelik çıkarma, yüz bulma.

ABSTRACT
REAL-TIME FACE RECOGNITION USING EXTENDED
CONVOLUTIONAL NEURAL NETWORKS

Mehmet Ali KUTLUGÜN

Ph. D. Computer Science and Engineering (%30 English)

Supervisor: Dr. Yahya ŞİRİN

April, 2023 - 98 +XV Pages

This study's subject is face recognition systems based on deep learning in real-time images. The research was carried out within the scope of convolutional neural networks used to obtain features in face recognition systems. Real-time images were analyzed in four stages: "face detection," "feature extraction," "vector transformation," and "classification and performance measurement." In face recognition systems, light direction, reflection, and emotional and physical changes in facial expression are known as the main factors that make recognition difficult. These elements, which can occur in every image compared to flowing image frames, lead to erroneous definitions. The background or scene changes detection method was proposed in the research, and the comparison of the features obtained from the convolutional neural networks was used together. This model detects blocks where the scene has not changed and tries to more precisely determine the comparison threshold value used in the classifier stage with a new value. Increasing the precision ratio between unchanging scene blocks allows for fewer comparisons between instances in the database. The model proposed in the experimental study reached 99,25% accuracy and 99,28% F-1 values compared to the original deep metric learning model. Experimental results show that even if there are differences in facial images of the same person in unchanging scenes, false recognition can be minimized as the compared sample area is narrowed.

Keywords: Deep metric learning, image processing, face detection, feature extraction, real-time face recognition.

İÇİNDEKİLER

TEZ ONAYI.....	i
BİLİMSEL ETİK BİLDİRİMİ.....	ii
ÖN SÖZ.....	iii
ÖZET.....	iv
ABSTRACT.....	v
İÇİNDEKİLER.....	vi
TABLO LİSTESİ.....	ix
ŞEKİLLER LİSTESİ.....	x
KISALTMALAR.....	xiv

BİRİNCİ BÖLÜM

GİRİŞ	1
1.1. Problem	2
1.2. Hipotez	3
1.3. Literatüre Katkı Hedefi	4
1.4. Tezin Amacı	4
1.5. Tezin Kapsamı.....	5
1.6. Varsayımlar	5

İKİNCİ BÖLÜM

YÜZ TANIMA.....	7
2.1. Genel Bilgiler	7
2.2. Yüz Tanıma Süreci.....	10
2.2.1. Yüz Tespiti	13
2.2.2. Yüz Hizalama/Düzleştirme ve İdeal Yüz Verisi Elde Etme	18
2.2.3. Yüze Ait Özellik Vektörü Elde Etme.....	19
2.3. Yüz Tanıma Sistemleri Üzerine Yapılan Araştırmalar	21
2.4. Yüz Tanıma Sistemlerinde Karşılaşılan Bazı Problemler	24

ÜÇÜNCÜ BÖLÜM

DERİN ÖĞRENME.....	27
3.1. Derin Öğrenme Nedir?	27

3.2.	Evrişimli Sinir Ağları	28
3.3.	Derin Öğrenme Ağları İşlemler ve Katmanları.....	30
3.3.1.	Giriş Katmanı	31
3.3.2.	Evrişim (Konvolüsyon) Katmanı	31
3.3.3.	Aktivasyon Katmanı.....	32
3.3.4.	Havuzlama (Pooling) Katmanı.....	33
3.3.5.	Tam Bağlı (Full-Connected) Katman.....	34
3.3.6.	Dropout Katmanı.....	35
3.3.7.	Sınıflandırma ve Yumuşatma (Softmax) Katmanı.....	35
3.3.8.	Normalizasyon (Normalization) İşlemi.....	36
3.4.	Derin Öğrenme Temel Algoritmaları	36
3.4.1.	LeNet.....	36
3.4.2.	AlexNet	37
3.4.3.	VggNet	38
3.4.4.	GoogleNet	39
3.4.5.	ResNet	41
3.5.	Veri Artırma	44
3.5.1.	Veri Artırma Üzerine Yapılan Araştırmalar.....	46
3.6.	Boyut Azaltma.....	47
3.6.1.	PCA	48
3.6.2.	LDA.....	49
3.6.3.	T-SNE.....	50

DÖRDÜNCÜ BÖLÜM

DENEYSEL ÇALIŞMA ve UYGULAMA GERÇEKLEME	51
4.1. Deneysel Çalışmanın Amacı	51
4.2. Uygulama-1: Veri Artırma Yöntemleri ile Aşırı Uyum Probleminin Ele Alınması	52
4.2.1. Veri Kümesi (Dataset).....	52
4.2.2. Veri Ön işleme	52
4.2.3. Uygulama Gerçekleme	52
4.2.4. Uygulama Değerlendirme	54
4.3. Uygulama-2: Yüz Tanıma Sistemlerinde Kademeli Benzerlik Oranları Yöntemi ile Eğitim Veritabanının Artırılması	56

4.3.1. Amaç	56
4.3.2. Model ve Mimari.....	57
4.3.3. Kademeli Benzerlik Oranları (Gradual Similarity Ratios, GSR).....	58
4.3.4. Uygulama	60
4.3.5. Uygulama Sonuçları ve Değerlendirme	64
4.4. Uygulama-3: Arka Plan değişimlerinin Gerçek Zamanlı Görüntüler Üzerinde Bir Parametre Olarak Kullanılması.....	68
4.4.1. Sahne Değişikliklerini Algılama Yöntemleri.....	69
4.4.2. Model	70
4.4.3. Sahne Değişim Göstergesi Bloğu (Scene Change Indicator, SCI).....	71
4.4.4. Veri Kümesi (Dataset).....	72
4.4.5. Yöntem	75
4.4.6. Uygulamanın Gerçekleştirimi	75
4.4.7. Uygulamanın Analiz Edilmesi	82
4.4.8. Değişken Hassasiyet Oranları Analizi.....	83
BEŞİNCİ BÖLÜM	
SONUÇ.....	85
5.1. Sınırlamalar ve Öneriler.....	85
KAYNAKÇA.....	87
ÖZGEÇMİŞ.....	97

TABLULAR LİSTESİ

Tablo 2.1. Boyutları farklı özellik vektörlerinin başarımlarını karşılaştırması.....	20
Tablo 2.2. LFW veri seti üzerinde uygulanan bazı mimarilerin karşılaştırması... ..	22
Tablo 2.3. Ülkemizde yazılan yüz tanıma konusu ile ilgili bazı doktora tezleri....	23
Tablo 4.1. Veri artırma başarımlarını değerlendirme tablosu.....	54
Tablo 4.2. Test videosunun özellikleri.....	63
Tablo 4.3. Farklı veri artırma yöntemlerinin 3 seviyeli performans tablosu.....	64
Tablo 4.4. Boyut küçültme etkileri tablosu.....	66
Tablo 4.5. Veri kümesinin özellikleri.....	74
Tablo 4.6. Donanım ve işletim sistemi bilgi tablosu.....	76
Tablo 4.7. Önerilen modelin performans tablosu.....	79
Tablo 4.8. CFM modelinin performans tablosu.....	80
Tablo 4.9. Dlib+K-nn ve SCI modellerinin performans tablosu.....	80

ŞEKİLLER LİSTESİ

Şekil 1.1. Poz açısı farklılıklarının sebep olduğu hatalı tanıma örneği.....	3
Şekil 2.1. Yüze ait metrik veriler.....	8
Şekil 2.2. Prosopagnozi örneği.....	9
Şekil 2.3. Google glass.....	10
Şekil 2.4. Temel yüz tanıma süreci.....	11
Şekil 2.5. Yüz tanıma akış diyagramı.....	12
Şekil 2.6. Derin öğrenme temelli ayrıntılı yüz tanıma süreci.....	12
Şekil 2.7. Gradyanın yönü ve büyüklüğü.....	14
Şekil 2.8. Her bir 8 x 8 piksellik görüntü bloğu için gradyanların çıkarılması.....	15
Şekil 2.9. Yüz görüntüsünün ana özelliklerinin HOG yöntemiyle temsil edilmesi...	15
Şekil 2.10. Haar benzeri dikdörtgensel öznitelikler.....	16
Şekil 2.11. Yüz bölgelerinin tespit edilmesi için öznitelik tarama.....	16
Şekil 2.12. Dlib yöntemi ile yüz bölgelerinin tespiti.....	17
Şekil 2.13. HOG-CNN tabanlı yüz tespit karşılaştırması.....	17
Şekil 2.14. 68 adet yüz işaretinin tespit edilmesi.....	18
Şekil 2.15. Yüz hizalama ve ideal yüz görüntüsü elde etme.....	19
Şekil 2.16. Üçlü kayıp modeli.....	19
Şekil 2.17. Yüz tespitini zorlaştıran örnek durumlar.....	25
Şekil 2.18. Yüz bulunmayan bölgede yüz tespiti durumu.....	25
Şekil 2.19. Yüz işaret noktalarının elde edilememesi durumu.....	26
Şekil 2.20. Uygun açıda olmayan kalabalık ortam yüz görüntüleri.....	26
Şekil 3.1. Yapay sinir ağı yapısı.....	27
Şekil 3.2. Derin sinir ağı yapısı.....	28

Şekil 3.3. ESA mimarisi.....	29
Şekil 3.4. ESA temel katmanları.....	30
Şekil 3.5. ESA temel özelliklerin ortaya çıkarılması.....	31
Şekil 3.6. Evrişim işlemi.....	32
Şekil 3.7. ESA katmanlarından elde edilen özellikler.....	32
Şekil 3.8. ReLU aktivasyon fonksiyonu.....	33
Şekil 3.9. Maksimum havuzlama örneği.....	34
Şekil 3.10. Tam-bağlı katman yapısı.....	34
Şekil 3.11. Dropout işlemi.....	35
Şekil 3.12. LeNet mimarisi.....	37
Şekil 3.13. AlexNet mimarisi.....	38
Şekil 3.14. VggNet mimarisi.....	39
Şekil 3.15. GoogleNet inception modülü örneği.....	40
Şekil 3.16. GoogleNet mimarisi.....	40
Şekil 3.17. Residual blok fonksiyonu.....	41
Şekil 3.18. ResNet mimarisi.....	43
Şekil 3.19. Bazı ESA mimarilerinin başarımlarını karşılaştırması.....	44
Şekil 3.20. Artırılmış görüntü örneği.....	45
Şekil 3.21. PCA vektörleri.....	49
Şekil 3.22. LDA vektör ve sınıfları.....	50
Şekil 4.1. Kayan çerçevelerde yanlış pozitif tanıma.....	51
Şekil 4.2. Süreç Blok Diyagramı.....	53
Şekil 4.3. Veri artırma ile elde edilen görüntüler.....	53
Şekil 4.4. Veri artırma uygulanmadan önce elde edilen eğitim grafiği.....	54
Şekil 4.5. 6300 adet veri artırma uygulanarak elde edilen eğitim grafiği.....	55
Şekil 4.6. 9450 adet veri artırma uygulanarak elde edilen eğitim grafiği.....	55

Şekil 4.7. İdeal sınıf ayrımı ile veri dağılımı.....	57
Şekil 4.8. Mimari.....	58
Şekil 4.9. GSR eylemi ile eşik ilişkisi.....	58
Şekil 4.10. Kademeli benzerlik oranlarına göre dağıtılan veriler.....	60
Şekil 4.11. Geleneksel yöntemler uygulanarak artırılmış görüntüler.....	61
Şekil 4.12. Farklı sınıflar arasında uygulanan MixUp görüntüleri.....	61
Şekil 4.13. Aynı sınıfa uygulanan MixUp görüntüleri.....	61
Şekil 4.14. FashionMNIST veri kümesi için MixUp örneği.....	62
Şekil 4.15. Farklı sınıflar arasında uygulanan CutMix yöntemi ile elde edilen görüntüler.....	62
Şekil 4.16. GSR ile veri artırma test videosunda yüz tanıma performansının karşılaştırılması.	63
Şekil 4.17. Rastgele artırılmış eğitim veri kümesi dağılımı.....	65
Şekil 4.18. GSR uygulanan veri kümesi dağılımı.....	65
Şekil 4.19. Farklı yöntemlerin başarımların karşılaştırılması.....	66
Şekil 4.20. Sahne değişim göstergesi ile gerçek zamanlı yüz tanıma süreci.....	68
Şekil 4.21. Arka arkaya üç kez aynı değişim özelliğini gösteren piksellerin tespit edilmesi.....	69
Şekil 4.22. Zincirleme sahne geçiş grafiği örneği.....	70
Şekil 4.23. Önerilen modelin mimarisi.....	70
Şekil 4.24. Veri Kümesi örneği.....	73
Şekil 4.25. Gerçek zamanlı video görüntülerde farklı eşik değerlerinin karşılaştırılması.....	76
Şekil 4.26. Video No-3 için farklı model eşiklerinin karşılaştırılması.....	77
Şekil 4.27. Video No-7 için farklı model eşiklerinin karşılaştırılması.....	77
Şekil 4.28. Farklı modellerin FP değişim grafiği.....	81
Şekil 4.29. Farklı modellerin ortalama performans karşılaştırılması.....	81

Şekil 4.30. Sabit ve aşırı hassas oran ile elde edilen görüntüler.....	82
Şekil 4.31. Uzak çekim, değişken hassasiyet oranı ile elde edilen görüntüler.....	83
Şekil 4.32. Eşik hassasiyet örnekleme.....	84



KISALTMALAR

API : Application Programming Interface : Uygulama Programlama Ara yüzü

CFSM :Continuous Frames Skipping Mechanism (Sürekli Çerçeve Atlama Mekanizması)

CNN : Convolutional Neural Networks (Evrışimli Sinir Ağları)

CPU : Central Process Unit (Merkezi İşlem Birimi)

DL : Deep Learning (Derin Öğrenme)

DÖ : Derin Öğrenme

ESA : Evrışimli Sinir Ağları

FN : False Negative : Yanlış Negatif

FP : False Positive : Yanlış Pozitif (Hatalı Onaylanmış)

FPS : Frames per Second (Saniyedeki Çerçeve Sayısı)

GPU : Grafical Processing Unit (Grafik İşlem Birimi)

GSR : Gradual Similarity Ratios (Kademeli Benzerlik Oranları)

HOG : Histogram of Oriented Gradient (Yönlendirilmiş Gradyanların Histogramı)

ILSVRC : ImageNet Large Scale Visual Recognition Challenge (ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesi)

k-NN : k-Nearest Neighbour – k-En Yakın Komşu

LDA : Linear Discriminant Analysis (Doğrusal Diskriminant Analizi)

LFW : Labeled Faces in the Wild (Vahşi Doğada Etiketlenmiş Yüzler)

MLP : Multi-Layer Perceptron (Çok Katmanlı Algılayıcılar)

MNIST : Modified National Institute of Standards and Technology (Deđiřtirilmiř Ulusal Standartlar ve Teknoloji Enstitüsü)

PCA : Principal Component Analysis (Temel Bileřen Analizi)

ReLU : Rectified Linear Unit (Dođrultulmuř Lineer Birim)

SCI : Scene Change Indicator (Sahne Deđiřim Gōstergesi)

TBA : Temel Bileřen Analizi

TP : True Positive : Dođru Pozitif (Dođru Onaylanmıř)

t-SNE : t-Distributed Stochastic Neighbor Embedding (t-Dađıtılmıř Stokastik Komřu Yerleřtirme)

YGH : Yōnlendirilmiř Gradyanların Histogramı

YSA : Yapay Sinir Ađları

BİRİNCİ BÖLÜM

GİRİŞ

Derin öğrenme (DÖ), çoklu doğrusal olmayan dönüşümlerden oluşan model mimarileri ile verilerin üst düzeydeki soyutlamalarını modellemeye çalışan bir dizi makine öğrenmesi (MÖ) algoritmalarıdır. DÖ sistemleri, işlemci gücünün artması ve grafik işlemcilerdeki geliştirmelerle birlikte büyük veri analizi, ses tanıma, doğal dil işleme, görüntü sınıflandırma, yaya algılama, jenerik görsel tanıma ve yüz tanıma gibi birçok alanda kullanılmaktadır (Bilgiç vd., 2017; Kutlugün ve Şirin, (2018a; 2020)). DÖ, verinin temsili ile öğrenme temeline dayanmaktadır. Örnek olarak, görüntü işlemede piksellere ait yoğunluk değerleri ilgili görüntünün bazı temsili özellikleri olarak düşünülebilir. Buna benzer başka bazı özellikler veriyi çok daha iyi temsil edebilir (Kutlugün vd., 2019). Bunların ortaya çıkartılmasında DÖ yöntemleri, elle çıkarılan özelliklerden ziyade veri içerisindeki en iyi temsilleri ortaya çıkarmak için daha etkin algoritmalar sunarak avantaj sağlamaktadır (Şeker vd., 2017; Song ve Lee, 2013).

Yüz Tanıma Sistemleri, biyometrik turnike geçiş, güvenlik ve gözetim sistemleri, havaalanları, polis kontrolleri, bankalar gibi alanlarda hem kontrol amaçlı hem de suç oluşumuna engel olmak amacıyla veya aranan suçluların tespitinde yaygın olarak kullanılmaktadır (Moghaddam, 1999). Üstün teknoloji ürünü yüz tanıma sistemleri güvenliğin önem arz ettiği bölgelerde sisteme bağlı güçlü bir yapı sunarak güvenliği daha üst seviyelere çıkarmaktadır (Sağlam, 2008). Yüz tanıma sistemlerinin temeli, yüze ait özniteliklerin çıkarılması ve bu özniteliklerden faydalanarak karşılaştırma yapılması esasına dayanmaktadır. DÖ modelleri, incelenen veriye ait özniteliklerin çıkarılmasında önemli başarılar elde eden bir yaklaşım sunmaktadır. Yüz tanıma derin öğrenme, özniteliklerin öğrenilmesinde kullanıldığı gibi yüz doğrulama ve benzerlik oranının tespit edilmesinde de kullanılmaktadır (Özcan, 2014).

Yüz tespiti ve yüz tanıma uygulamaları akademik araştırma çevreleri kadar ticari uygulama geliştiren şirketler tarafından da oldukça ilgi gören sistemler haline gelmiştir. Özellikle bu sistemlerin gerçek-zamanlı olarak verimli bir biçimde uygulanabilmesi halen zorlayıcı bir alan olarak güncelliğini korumaktadır. Son zamanlarda derin öğrenme ve bilgisayarlı görü çalışmaları yüz tespiti ve tanıma

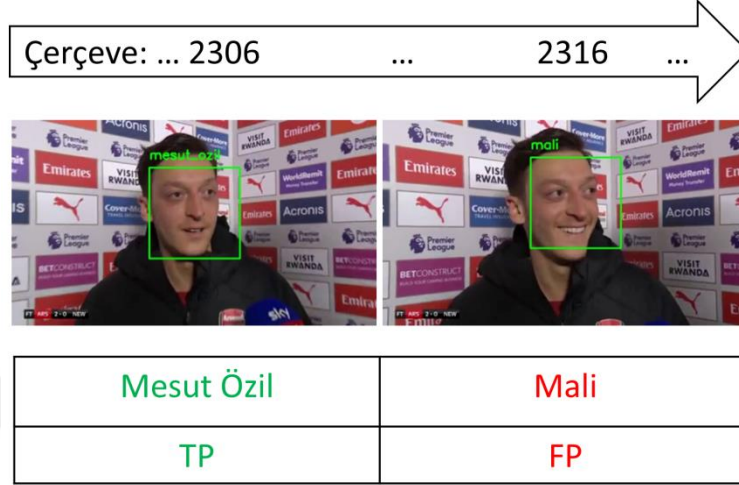
çalışmalarına katkı sağlayan alanlar olarak öne çıkmaktadır. Yüz bulma ve tanıma sistemleri, farklı alanlarda uzman olan araştırmacılar tarafından gösterilen ilgi sayesinde gündelik ve sanal hayatta önemli bir yer edinmektedir. Yüz bilgilerinin yeniden kazanımı, insan-bilgisayar ara yüzü geliştirme, multimedya haberleşmesi, yapay yeni yüzlerin elde edilmesi ve içerik tabanlı yüz veri depolarının yönetimi gibi birçok sistem gelişmiş yüz tanıma teknolojilerine ihtiyaç duymaktadır. Bu sistemler uygun koşullarda verimli sonuçlar verse de, şartlardaki küçük farklılıklar yüz bulma ve tanıma işlemlerini olumsuz etkileyebilmektedir (Sütçüler, 2006).

1.1 Problem

Bir hareketli görüntü dizisinin kalitesi, görüntülerin uzamsal ve zamansal çözünürlüğü dâhil olmak üzere bir dizi faktöre bağlıdır. Modern dijital sinema endüstrisi, teknik olarak daha yüksek kalitede çözünürlük ve kare hızları sunma yeteneğine sahiptir (Allison vd., 2020). Yaklaşık 100 yıl önce film endüstrisi, kısmen istikrarlı görsel-işitsel etkileşimi kolaylaştırmak için sinema kare hızlarını saniyede 24 kare (fps) olarak standartlaştırmıştır (Watson, 2013). Saniyede 24 kare hızı ile oluşturulan görüntü kümeleri insan algısına uygun oran olarak kabul görmektedir. Çok özel bir amaç için oluşturulmadıkça bu kayan pencerelerin sürekli değişmesi beklenmez. Yani, izlenen bir görüntü içerisinde 1sn.'lik bir zaman dilimi içinde geçen 24 çerçevenin her birinin birbirinden çok farklı olması durumunun oluşması istenmez.

Gerçek zamanlı olarak hareket eden görüntü çerçeveleri, zaman zaman sahne değişimlerine, ortam değişikliklerine uğrayabilir. Bu durum genel bütünlüğün bozulduğu anlamını taşımamaktadır. Gerçek zamanlı yüz tanıma sistemlerinde de akan görüntü bir anda değişikliklere uğramadan belli bir zaman periyodunda hareket etmektedir. Ancak buna rağmen, görüntü kümesi içinde uygun akış devam ederken bazı bozulmalar veya hatalı tanımlamalar meydana gelebilmektedir. Yüz tanıma sistemlerinde ışığın yönü, yansımaları, yüz ifadesindeki duygusal ve fiziksel değişiklikler tanımayı zorlaştıran temel etkenler olarak karşımıza çıkabilmektedir (Kutlugün vd., 2019). Bu durumda sistem, doğru tanınması gereken kişi yerine farklı bir kişi olarak ezberleme veya aşırı uydurma (overfitting) eğilimine girebilmektedir. Aşağıdaki Şekil 1.1'de kayan görüntü içinde ortam ve kişi değişikliği olmamasına rağmen bazı çerçevelerde oluşan poz açısı değişikliği yanlış tanımlamaya sebep

olmaktadır. Model, genel olarak kişiyi doğru tanımlasa da poz farklılıkları durumunda zaman zaman kendisine en fazla benzeyen kişiye benzetmeye çalışma veya başka bir ifadeyle veri kümesindeki en fazla benzerlik gösteren özellik vektörüne aşırı uyum sağlama eğilimine girebilmektedir.



Şekil 1.1. Poz açısı farklılıklarının sebep olduğu hatalı tanıma örneği.

1.2 Hipotez

Yüz tanıma sistemlerinde meydana gelebilecek tanımayı olumsuz etkileyen etmenler birçok farklı sebepten kaynaklanabilmektedir. Bunlar, yüz tespitinde karşılaşılan hatalı durumlar, eğitimi olumsuz etkileyen parametreler ile oluşturulan eğitilmiş bir model, en uygun öznelik elde edilmesini zorlaştıran sebepler gibi sürecin birçok aşamasında karşımıza çıkabilmektedir.

Kayan pencereler üzerinde, ortam ve o an görüntüde bulunan kişi değişmese bile, hatalı tanımlamaya sebep olacak durumların azaltılması genel başarıyı artıracaktır. Ayrıca, görüntü akışının bozulmadan bir bütünlük içerisinde hareket etmesi sağlanmış olacaktır. Bu hataların azaltılması için kullanılacak yöntemlerden biri, artırılmış veri kümeleri ile çok iyi eğitilmiş bir model elde etmek olabilir. Bu durumda elde edilmesi hedeflenen veri kümesinin boyutu veya tanınması istenen kişilerden kaç farklı görüntü olması gerektiği, hangi farklı veri artırma metodunun daha etkili olduğu, eğitim süresi gibi birçok durumun birlikte değerlendirilmesi gerekmektedir.

1.3 Literatüre Katkı Hedefi

Bu tez kapsamında, gerçek zamanlı yüz tanıma sürecinin bir bütünlük içerisinde ele alınmasının gerektiği düşüncesi ile hareket edilmektedir. Bu amaçla, tanıma performansına katkı sağlayacak ve olası aşırı uydurma, ezberleme ve hatalı onaylanmış (false positive) durumlarının ayıklanabilmesi için yeni parametrelerin sisteme dâhil edilmesine çalışılmıştır. Bu çalışma ile;

- Az sayıda örneğe sahip bir veri kümesi üzerinde açısız döndürme, kaydırma, kırpma, yakınlaştırma, çevirme gibi filtreleri uygulayarak kaynak veri kümesinin bütünlüğünün bozulmadan artırılması ve kullanılan hangi filtrelerin başarıma daha fazla olumlu yönde etki ettiğinin tespit edilmeye çalışılması,
- Sınırlı sayıda örnek içeren veri kümelerinde gelişigüzel veri artırmak yerine, kullanılan sınıflandırıcı eşik değerinin de dikkate alınarak filtreleme yapılması ve gereksiz artırılmış verilerin kaldırılması,
- Öznitelik vektör boyutunun varsayılan değerden daha düşük oranlarda işlenebilmesi için boyut azaltma uygulanması,
- Önerilen arka plan değişim göstergesi ile sahnenin değişikliğe uğrayıp uğramadığı bilgisini parametre olarak gerçek zamanlı yüz tanıma bütünlüğünün korunmaya çalışılması,

şeklinde literatüre katkı sağlama hedeflenmektedir.

1.4 Tezin Amacı

Bu tez çalışmasının amacı, anlık görüntü akışına en uygun modelin gerçek zamanlı yüz tanıma sistemlerinde yüksek başarımla çalışmasıdır. Bu anlık değişen karelerde karşılaştırma biçimine göre değişen kayıplar meydana gelebilmektedir. Belirlenen en uygun değer hassasiyet oranı ile kişinin doğru tanınıp tanınmadığı ölçülmektedir. Ancak başlangıçta belirlenen bu değer en ideal şekilde seçilmemesi yanlış karşılaştırma ile sonuçlanmaktadır.

Bu çalışmada, yüz tanıma süreci ve gerçek zamanlı yüz tanıma sistemlerinde karşılaşılan zorluklar açıklanarak, arka plan değişikliklerinden elde edilen yeni bir parametre yardımıyla karşılaştırma hassasiyet oranının dinamik olarak değişmesini sağlayacak bir model önermektedir. Bu modele göre, odaklanan görüntü değişmedikçe sabit bir değer ile karşılaştırma yapılırken, arka plan görüntü değişimlerinde hassasiyet oranı otomatik olarak yeniden belirlenmektedir. Böylece, aynı kişiye ait görüntülerde poz açısı veya ışık şiddeti gibi etmenler değişse de odak noktası daha az değişime uğrayarak yanlış kişi tanımlamasının önüne geçilmesi amaçlanmıştır. Özet olarak, Yanlış Pozitif tespit edilen karelerin sayısının azaltılması hedeflenmiştir.

1.5 Tezin Kapsamı

Bu tez kapsamında gerçek zamanlı akan video görüntülerinde yüz bölgeleri içeren görüntülerin derin metrik öğrenme yöntemleri ile elde edilen öznitelikler yardımıyla kişi tanımlaması yapılmaktadır. Tanınması istenen kişilerden az sayıda örnek veri kümesi ile eğitilmek suretiyle veri tabanı oluşturulmuştur. Bu işlemler sırasında derin öğrenme modellerinden faydalanılmış ve az sayıda eğitim verisi ile en yüksek başarımlar elde edilmeye çalışılmıştır. Önerilen model, daha önce eğitilen verilerden farklı olarak seçilen gerçek zamanlı görüntüler üzerinde başarımlar değerlendirilmiştir. Bu gerçek zamanlı görüntüler, çok kalabalık gruplar içermeyen ve belli sahnelerde akışın değiştiği video görüntülerinden seçilmiştir.

1.6 Varsayımlar

Tez genelinde genel kabul görmüş başarımları yüksek derin metrik öğrenme yöntemlerinin uygulandığı modellerden dlib modeli tercih edilmiştir. Sınıflandırıcı aşamasında ise k-en yakın komşu algoritması ile yüz tespiti gerçekleştirilmiştir. Ancak sadece bu kapsama bağlı kalmadan gerçek zamanlı görüntülerde başarımları olumsuz etkileyen unsurlara dikkat çekilerek bu alanlara özel çözüm önerileri sunulmuştur. Ayrıca yapılan uygulama çalışmalarında çok yüksek performanslı bilgisayarlar yerine, standart özelliklerde bilgisayar donanımları tercih edilmiştir. Paralel işlemler veya GPU donanımlı cihazlar kullanılmadan özellikle başarılı tanıma performansını artıracak çözümler üzerinde durulmuştur. Bunun en büyük sebebi zaten ön-eğitilmiş bir model tercih edilmesinden dolayıdır. Yani, modeli en baştan yeniden eğitmeye ihtiyaç duyulmamaktadır. Bu alanda özellik çıkarımı açısından

önemli başarımları sağlayan bir model kullanıldığı için varsayılan olarak bu modelin değerleri dikkate alınarak geliştirmeler yapılmıştır.



İKİNCİ BÖLÜM

YÜZ TANIMA

2.1 Genel Bilgiler

Günümüzde insan bilgisayar etkileşimli sistemlerde, biyometrik özelliklerin (yüz, ses) etkin bir faktör olarak kullanıldığı görülmektedir. (Lu, 2003). Geçtiğimiz uzun yıllar boyunca yüz tanıma konusunda yapılan araştırmalar ve gerçekleştirilen uygulamalar, yüze ait biyometrik özelliklerin en fazla tercih edilen sistemler olduğunu düşündürmektedir (Yazıcı, 2008).

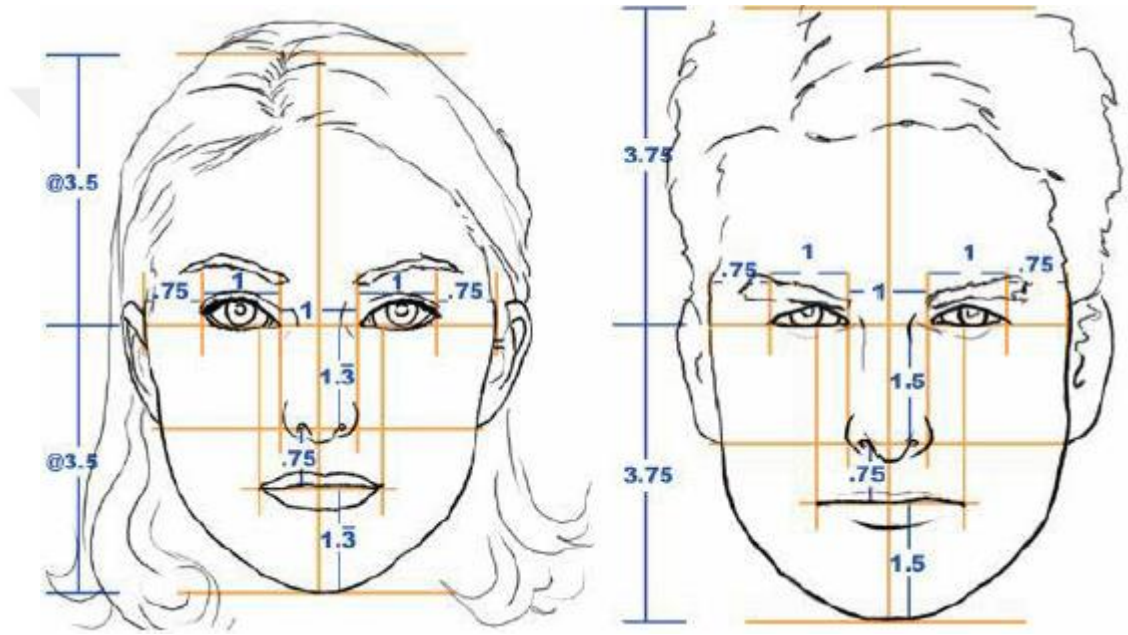
Yüz tanıma problemleri özellikle görüntü işleme ve bilgisayarla görü gibi alanlarda güncel olarak araştırmacıların yoğun bir biçimde üzerinde çalıştıkları zor problemler arasında sayılmaktadır (Kahraman vd., 2005). Bu konu üzerinde uzun yıllar boyunca çalışılmasına ve birçok farklı uygulama geliştirilmesine rağmen, gerçek yaşamda karşılaşılan sorunlara tek başına yeterli gelecek başarı oranına halen ulaşamamıştır. Özellikle, artan güvenlik gereksinimleri gibi ihtiyaçlar sebebiyle biyometrik yüz tanıma alanına ilgi giderek artmaktadır. Görüntü işleme ve bilgisayarlı görü alanlarında gerçekleştirilen uygulamalar, akıllı kamera sistemleri, yapay zekâ, savunma sanayi ve robotik sistemler, tıbbi sistemler, sinyal işleme ve makine öğrenmesi gibi birçok farklı alana hizmet etmektedir (Subaşı, 2011).

Yüz tanıma sistemi, 1970'lerden beri araştırmacıların üzerinde çalıştığı biyometrik bir tekniktir. Yüz üzerinde kişilerin kimliklerini tanıma, başlangıçta iki boyutlu örüntü tanıma problemi olarak ele alınmıştır. İlk çalışmalarda, yüz üzerinde mevcut önemli bazı ayırt edici özelliklerin tespit edilmesi ve bu noktalar arasındaki mesafenin hesaplanarak veri tabanında bulunan görüntüler üzerinden sınıflandırma yapılması hedeflenmiştir (Özcan, 2014). Kabaca şimdiye kadar yapılan yüz tanıma çalışmalarının iki ana yöntem ile ele alındığı söylenebilir.

Geleneksel yöntemler, bir yüz üzerindeki özel geometrik özellikler yaklaşımına göre hareket etmekte ve bu şekilde bir görüntüyü sayısal değerlere dönüştürmeye çalışmaktadır. Bunların bazıları bir yüz şablonu ile eşleme yaparak fotometrik bir yaklaşım benimsemektedir. Geleneksel yöntemlere örnek olarak Fisher Vector algoritması gösterilebilir (Sufiyan, 2018).

Geleneksel yöntemler insan yüzünde belirlenmiş bazı değerleri dikkate almaktadır. İnsan yüzünde yer alan bazı değerler aşağıda sıralanmıştır (Kaplan, 2018). Şekil 2.1’de yüze ait metrik veriler temsili olarak gösterilmektedir.

- Yüzün eni ve boyu,
- Kaşlar ile dudak arası mesafe,
- Kaşlar ile çene arası mesafe,
- Ağız ve burun genişliği,
- Burun eni ve boyu,
- Göz ile kaşlar arası mesafe.



Şekil 2.1. Yüze ait metrik veriler

Kaynak: (Kaplan, 2018).

Modern yüz tanıma yöntemleri, DL tekniklerine dayanmaktadır. Çok büyük miktarda veri ve daha güçlü CPU ve GPU'ların varlığı nedeniyle, araştırmacılar birçok farklı ortam ve koşulda çalışabilen çok yüksek doğrulukta daha sağlam yüz tanıma sistemleri geliştirebilmişlerdir. Bu DL tabanlı yöntemlerin genel çalışma prensibi, dijital görüntüler biçimindeki verilerin bir DL algoritması ile beslenerek, zaman içinde verilerin bazı özelliklerini otomatik olarak öğrenebilmesidir. DL algoritması verilerden örüntüleri veya özellikleri yeterince öğrendikten sonra, daha önce

görmediği yeni bir dijital görüntüden örüntü çıkarıp veri kümesindeki örüntülerle karşılaştırma yapabilmektedir (Sufiyan, 2018).

Yüz tanıma sistemlerinin kullanıldığı özel çalışma alanlarından biri de Prosopagnozi olarak da bilinen yüz körlüğüdür. Yüz körlüğü, tanıdık kişilerin yüzlerinden ciddi şekilde tanınmaması anlamına gelir. İfade, saç modeli ve giyimdeki değişikliklere rağmen meslektaşlarımızı, arkadaşlarımızı ve ailemizi farklı yerlerde ve zamanlarda görsek bile kolayca tanırız. Ancak prosopagnozisi olan kişiler daha önce bir yüz gördüklerini anımsamazlar. Bu denekler bir yüzün yüz olduğunu anlarlar, ancak yüzü daha önce görüp görmediklerini ya da kimin yüzü olduğunu anlayamazlar. Bunu saç modeli, yürüyüş veya ses gibi diğer kimlik ipuçlarıyla eşleştirmeye çalışırlar, ancak bu ipuçları değişirse, kaçınılmaz olarak hata yapabilirler (Corrow vd., 2016). Yüz körlüğü problemi yaşayan birçok insan, yaş, cinsiyet, duygusal ifadeyi işleme veya bir kişinin bakış yönünü takip etme gibi zorluklarla karşılaşmaktadır. Bazı insanlar yüz körlüğü ile daha iyi başa çıkmakta ve hatta günlük yaşamda etkili bir şekilde çalışmasına yardımcı olmak için farklı araçlar kullanmaktadır. Ancak bir kısmı ise günlük aktivitelerde bile zorlanmaktadır. Bu sebeple, bazıları sosyal etkileşimlerden kaçınırken kişilerarası ilişkilerde zorluk yaşarlar (Yazıcı, 2008). Şekil 2.2’de bir prosopagnozi örneği gösterilmiştir.



Şekil 2.2. Prosopagnozi örneği

Kaynak: (Yazıcı, 2008).

Bu problem için sunulan çözümlerden biri yüz tanıma teknolojisi kullanılarak üretilen gözlüklerin kullanılmasıdır. Şekil 2.3'te Google tarafından tasarlanan bir yüz tanıma gözlüğü örneği gösterilmektedir (Leonard, 2014).



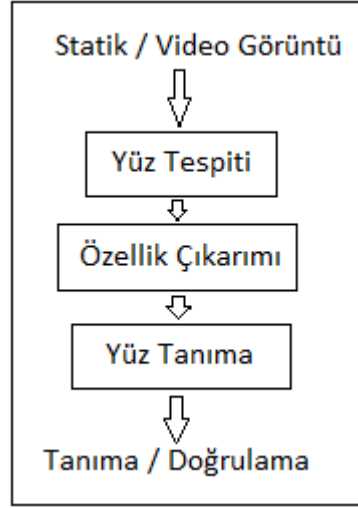
Şekil 2.3. Google glass

Kaynak: (Leonard, 2014).

2.2 Yüz Tanıma Süreci

İnsanların otomatik olarak çok hızlı bir biçimde algılayabildiği kişi tanıma veya kimlik belirleme işlemlerini bilgisayar sistemleri tarafından gerçekleştirmek zor bir görev olarak görülmektedir. Yüz tanıma sistemleri, günümüzde artık cep telefonlarına kadar birçok gelişmiş teknolojide faydalı birer araç olarak kullanılabilir. Bu teknolojiler birçok görüntü içerisinde farklı insanlara ait yüz özelliklerinin ayırt edilebilmesi için kullanılabileceği gibi, en benzer görüntülerin tespit edilerek gruplandırılması gibi amaçlarla da tercih edilmektedir (Atik, 2020; Soltanpour vd., 2017).

Yüz tanıma problemi genel olarak, statik veya bir video görüntüsünden sahne ortamı üzerindeki kişi veya kişilere ait yüz özelliklerinin, daha önceden eğitilmiş veri tabanında kaydedilen yüzlerin sayısal değerleri ile eşlenerek kimliğinin tespit edilmesi ya da doğrulanması şeklinde ifade edilebilir. Yüz tanıma sistemleri kabaca Şekil 2.4'teki gibi üç temel aşamadan oluşmaktadır. Bu adımlar; yüz içeren görüntü veya görüntülerin sahne ortamı üzerinde tespit edilmesi, tespit edilen yüzlerle ait ayırt edici özelliklerin ortaya çıkarılması ve son olarak kimliklendirme ya da doğrulama işlemlerinin yapılmasıdır (Yazıcı, 2008).



Şekil 2.4. Temel yüz tanıma süreci

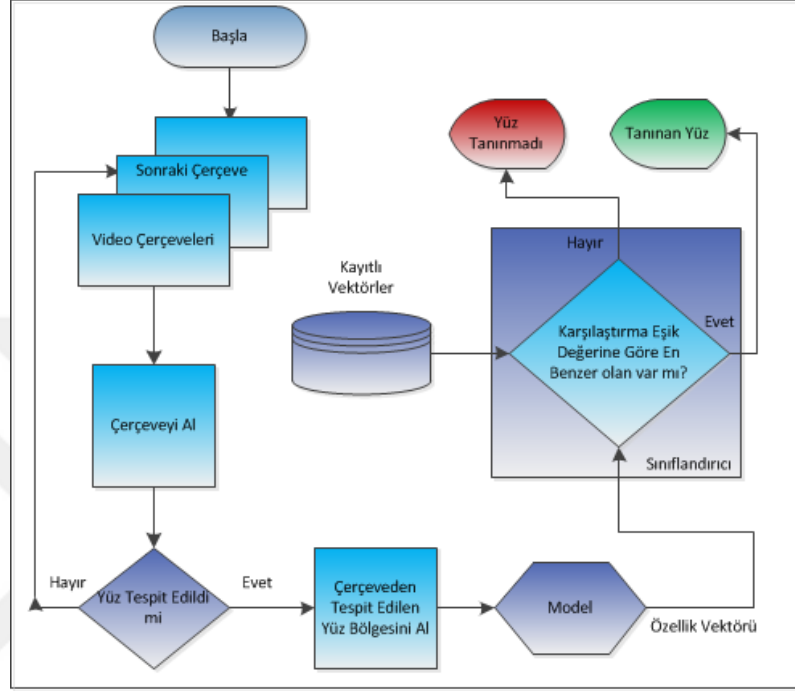
Kaynak: (Yazıcı, 2008).

Tanıma işleminde sorgulanan görüntünün veri tabanındaki kayıtlı etikete ait görüntülerden en çok hangisine benzediğinin bulunması, yüz tanıma sürecini bir yakınlık hesaplama problemine dönüştürmektedir. Bunun için sorgulanan resme ait ayırt edici özellikler, yüz veri tabanında kayıtlı diğer özellik vektörleri ile benzerlik karşılaştırması yapılarak bir yakınlık değeri elde edilmektedir. Bu yakınlık değeri sonucunda hedef görüntünün sınıfına karar verilmektedir. Bu modelin kalitesinin yüksek olması, aynı zamanda yüz tanıma veri tabanının içeriğine ve elde edilen yüz pozlarının kalitesine de doğrudan bağlıdır (Sütçüler, 2006).

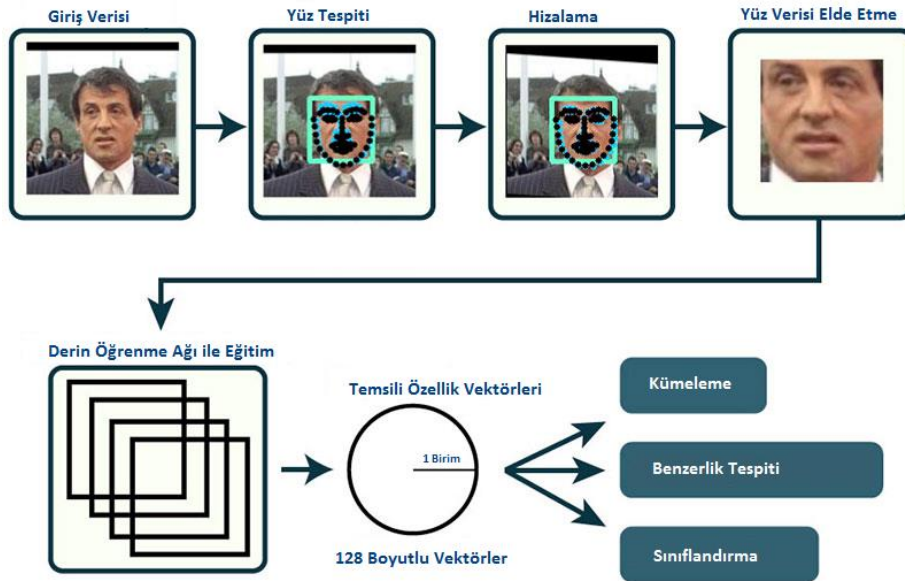
Yüz kimliklendirmenin en önemli aşaması, yüz görüntülerinden en ayırt edici eşlemeleri sağlayabilecek temsili özelliklerin ortaya çıkarılmasıdır. Böylece yüz tanıma sistemleri, insana ait biyometrik özelliklerin sayısal olarak işlenmesi ve kişiyi doğru etiketleyerek diğer yüz görüntülerinden oluşan bir kümede karşılaştırma, doğrulama veya sınıflandırma yapabilen sistemlerdir (Özkaya ve Sağıroğlu, 2008; Atik, 2020).

Bir yüz tanıma sistemi genelde, dijital görüntünün ilk olarak bir kamera veya başka herhangi bir görüntü toplama cihazı tarafından ele alınması ile başlamaktadır. Ardından alınan görüntü bazı ön işlemlerden geçirilerek yüz özelliklerinin çıkarıldığı modelin uygulanması aşamasına devam edilmektedir. Sonraki aşamada ise, elde edilen bu yüz özelliklerini sınıflandırmak için bir sınıflandırıcı kullanarak veri tabanında kayıtlı yüz özellikleri ile belli bir eşik değerine göre karşılaştırma

yapmaktadır. Büyük ölçüde veri tabanında kayıtlı bir eşleşme bulunması durumunda, sistem yüzü tanımlamakta, aksi halde bilinmeyen bir yüz olarak adlandırmaktadır (Sufiyan, 2018). Sürece uygun bir akış diyagramı Şekil 2.5'te gösterilmektedir. Şekil 2.6'da ise derin öğrenme temelli ayrıntılı yüz tanıma süreci görsel olarak verilmektedir.



Şekil 2.5. Yüz tanıma akış diyagramı.



Şekil 2.6. Derin öğrenme temelli ayrıntılı yüz tanıma süreci

Kaynak: (Amos vd., 2016).

Bu süreçlere ait ihtiyaç duyulan yukarıda adı geçen temel aşamalar aşağıda detaylıca açıklanmaktadır.

2.2.1 Yüz Tespiti

Sisteminin ilk aşaması olan yüz tespiti aşaması, görüntü içindeki bir veya birden çok yüz görüntülerinin sahne ortamı üzerinde tespit edilmesi anlamına gelmektedir. Temel olarak, görüntüler içinde yüz olan-olmayan bölgeler belirlenerek ayırım yapıldığından görüntü bölümlenme problemi olarak da düşünülebilir (Bozik, 2019).

Bu aşama, tüm yüz tanıma sistemini doğrudan etkilemektedir.

Literatürde, kullanılacak görüntüler içerisinde yüz ve yüze ait bileşenleri belirlemek için birçok yöntem önerilmiştir. Bunlardan en çok kullanılanlar aşağıda maddeler halinde sıralanmıştır.

- Yönlendirilmiş gradyanların histogramı
- Viola-Jones yöntemi
- YSA tabanlı yöntemler
- SVM tabanlı yöntemler
- Dlib yöntemi
- Yerel ikili desenler histogramları

Yüz tespitinde HOG algoritması ve Viola-Jones Algoritması sıklıkla kullanılan iki algoritmadır. Bu yöntemlerin her birinin daha etkili olduğu durumlar olmakla birlikte, uygulamadaki ihtiyaçların tam olarak belirlenmesi ve daha sonra hangi özelliğin etkisinin bu ihtiyaca cevap verdiğinin değerlendirilmesi gerekmektedir (Bozik, 2019).

Bu algoritmalarından bazılarının ait detaylı açıklamalar aşağıda sunulmuştur.

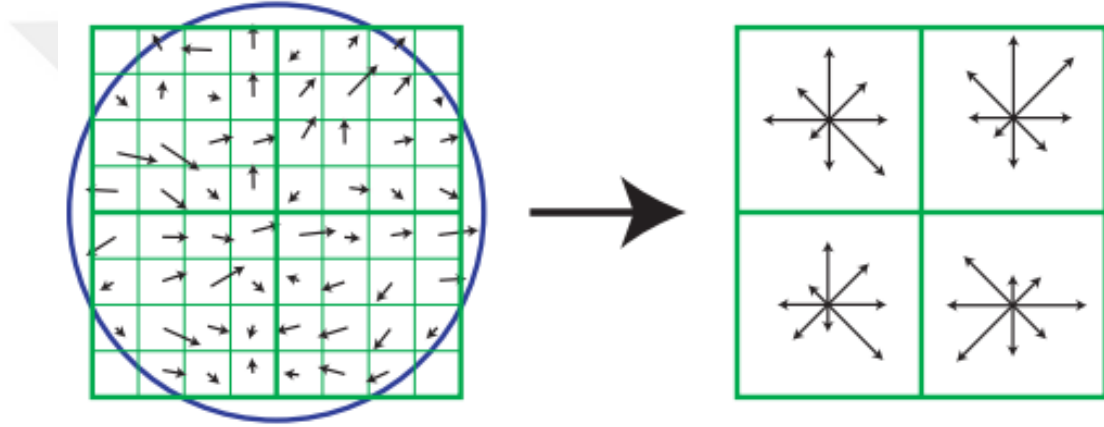
- Histogram of Oriented Gradients - HOG yöntemi

Görüntüdeki yoğunluk değişimlerini izleyip bunların yönlerine göre oluşturulan histogramlar sayesinde öznitelik elde etmeyi amaçlayan HOG yöntemi Dalal ve Triggs tarafından geliştirilmiştir (Dalal ve Triggs, 2005).

HOG algoritması, görüntünün yerel bir alanı içerisindeki gradyanları dikkate almaktadır. Bu, belli bölgedeki gradyanlara ait yönlerin sayılarını ifade etmektedir (Şahin, 2017).

Bir görüntü işleme kavramı olan gradyan, Şekil 2.7'deki gibi görüntüdeki yoğunluğu ve renk değişim yönünü belirtmektedir (Uçar, 2019). HOG yönteminin aşamaları aşağıdaki gibi sıralanabilir;

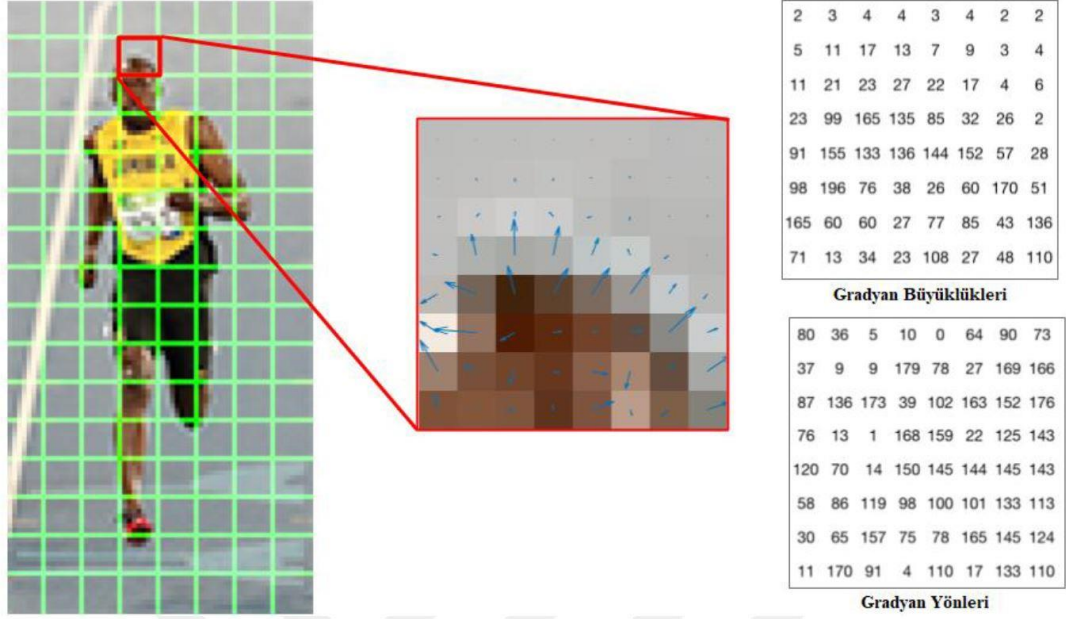
1. İmgeye ait gradyan alınması
2. İlgili lokasyonlarda yön histogramlarının belirlenmesi
3. Histogramların normalizasyonu (Karakaya vd., 2009).



Şekil 2.7. Gradyanın yönü ve büyüklüğü

Kaynak: (Lowe, 2004).

Şekil 2.8'de sol tarafta verilen görüntü ve bu görüntüye ait alt bloklar gösterilmektedir. Bu görüntü bloklarındaki gradyanların yönleri ortada gösterilmekte ve elde edilen yön matrisleri ise okların uzunlukları ile sağ tarafta gösterilmektedir. Gradyan yön matrisinde açılar $0-360^\circ$ yerine $0-180^\circ$ arasında ifade edilmektedir. Gradyanlar elde edildikten sonra her bir bloğa ait histogramlar çıkarılmaktadır. Histogramda ise $0^\circ, 20^\circ, 40^\circ, 60^\circ, 80^\circ, 100^\circ, 120^\circ, 140^\circ$ ve 160° olmak üzere 9 adet açı bulunmaktadır. Bir görüntünün gradyanları ışıklandırma değişimlerine duyarlıdır (Uçar, 2019).



Şekil 2.8. Her bir 8 x 8 piksellik görüntü bloğu için gradyanların çıkarılması

Kaynak: (Uçar, 2019).

Daha sonra normalize edilen bloklar, görüntünün yalnızca ana hatlarını bırakarak bir yüzü algılamak için kullanışlı olmayan birçok bilgiyi ortadan kaldırmaktadır. Şekil 2.9'da HOG yöntemi kullanılarak işlenen bir görüntü gösterilmektedir. Gradyanların yönü, görüntünün daha açık bölgelerinden daha koyu bölgelerine işaret etmektedir (Sufiyan, 2018).



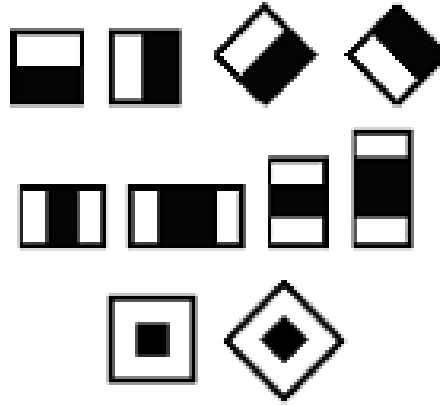
Şekil 2.9. Yüz görüntüsünün ana özelliklerinin HOG yöntemiyle temsil edilmesi

Kaynak: (Sufiyan, 2018).

- Viola-Jones Algoritması

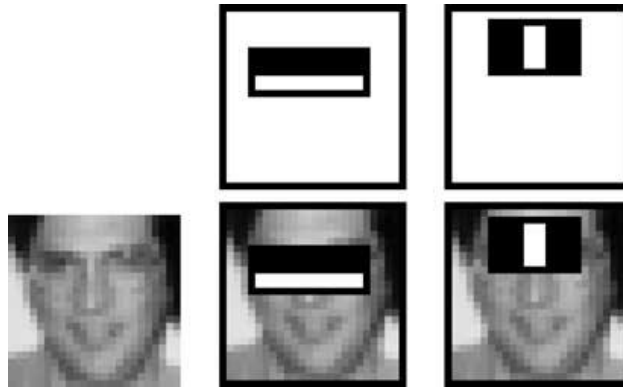
Literatürde Viola-Jones olarak bilinen algoritma yüz tespitinde oldukça yaygın kullanılarak başarılı sonuçlar vermektedir. Yüz tespitini gerçekleştirmek için Şekil 2.10'da gösterilen çerçeveler, görüntüler üzerinde Şekil 2.11'deki gibi hareket ettirilmektedir. Görüntüdeki siyah-beyaz alanlara denk gelen bölgelerdeki piksel değerleri toplanarak toplamların farkı ile özellikler belirlenmektedir. Çıkan sonuç değeri, belirlenen bir eşik değeri ile karşılaştırılır ve bu eşik değerden büyük olan durumlar nesnenin algılandığını, küçük olduğu durumlar ise nesnenin algılanmadığını ifade eder (Öznel, 2019).

İşlenen bir görüntüde Viola Jones algoritması HOG yöntemine göre daha hızlı olmasına rağmen, HOG yöntemi daha yüksek doğruluk ile çalışmaktadır (Bozik, 2019).



Şekil 2.10. Haar benzeri dikdörtgensel öznelikler

Kaynak: (Atasoy, 2016).

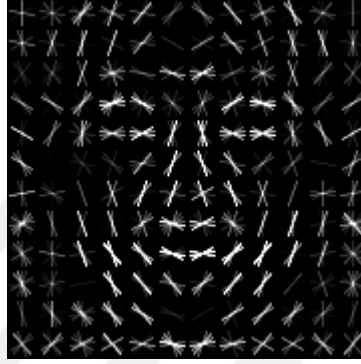


Şekil 2.11. Yüz bölgelerinin tespit edilmesi için öznelik tarama

Kaynak: (Viola, 2004).

- Dlib Yöntemi

Dlib kütüphanesi, Davis King tarafından geliştirilen ve üzerinde birçok derin öğrenme algoritmaları barındıran, C++ ve Python API'sine sahip olduğu için oldukça geniş bir kitleye hitap eden bir kütüphanedir (Ayata, 2020). HOG yönteminin geliştirilmiş sürümü olarak tanımlanan Dlib, temel olarak HOG özellik çıkarıcı ile evrişimli sinir ağlarını birleştirip daha iyi performans ve doğruluk sağlamak için üretilmiştir (King, 2009). Şekil 2.12'de Dlib yöntemi ile yüz bölgelerinin tespiti örneği gösterilmektedir.

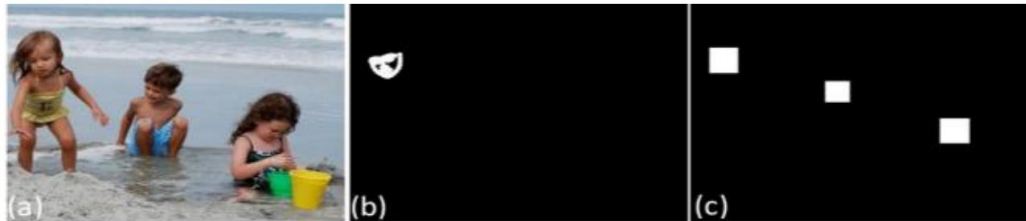


Şekil 2.12. Dlib yöntemi ile yüz bölgelerinin tespiti

Kaynak: (King, 2009).

- Yapay Sinir Ağı Tabanlı Algoritmalar (CNN temelli yüz tespiti)

Yapay sinir ağı tabanlı yüz tespiti algoritmaları, özellikle kötü pozlanmış yüzlerde, HOG tabanlı detektörden daha iyi performans göstermektedir (Kong vd., 2018). HOG tabanlı algoritmalar yüzün önden görüntüsünün daha belirgin olduğu durumlarda doğru sonuçlar verirken, Şekil 2.13'te görüldüğü gibi CNN tabanlı algoritmalar farklı poz açılarında da doğru bölgeleri işaret edebilmektedir.



Şekil 2.13. HOG-CNN tabanlı yüz tespit karşılaştırması a)Alınan görüntü b) HOG algoritması ile tespit c) CNN tabanlı tespit

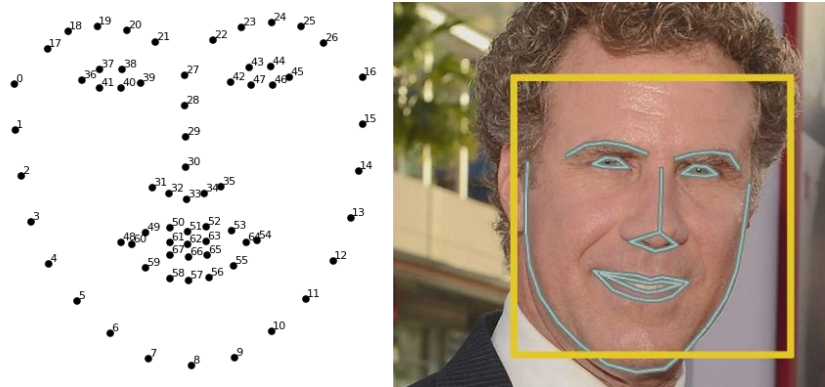
Kaynak: (Kong vd., 2018).

Bunun yanında CNN tabanlı algoritmalar HOG algoritmasına göre daha yavaş sonuç üretmektedir. Bunun için GPU ile çalışan güçlü donanımlara sahip cihazların tercih edilmesi önerilmektedir (Rosebrock, 2018).

2.2.2 Yüz Hizalama/Düzleştirme ve İdeal Yüz Verisi Elde Etme

Yüz görüntüleri her zaman aynı konumda veya yönde olmayabilir ve bu durumda yüz tanıma sistemi, farklı yön ve konumlardaki yüzleri farklı kişiler olarak görebilmektedir. Yüz görüntüleri elde edildikten sonra, yüz hizalama ile daha düz bir görüntü elde etmek ve böylece başarıyı artırmak hedeflenmektedir. Gözlerin ve dudakların orijinal yönleri ne olursa olsun, aynı örnek konumunda olması için tüm görüntülerin düzleştirilmesi gerekmektedir (Sufiyan, 2018). Bu işlem için birçok farklı yöntem mevcuttur.

Hizalama yöntemlerinden 2014 yılında Vahid Kazemi ve Josephine Sullivan tarafından geliştirilen yüz dönüşüm noktası algoritması çokça tercih edilmektedir (Kazemi ve Sullivan, 2014). Hizalama işlemine bir görüntüdeki yüzden, çoğunlukla her iki gözün dış kenarları, çenenin üst kısmı, her iki kaşın iç kenarları, burun vb. yüz içinden 68 tane belirli noktanın Şekil 2.14'teki gibi elde edilmesi ile başlanmaktadır (Geitgey, 2017). Bu işlem için kullanılan daha farklı yöntemler mevcuttur. 68 adetten daha az veya yüz içinde sadece göz bölgeleri ile tespit ve düzleştirme çalışmaları bunlara örnek olarak gösterilebilir.

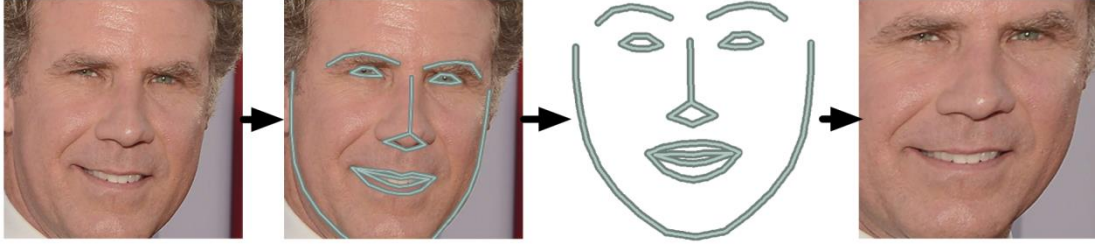


Şekil 2.14. 68 adet yüz işaretinin tespit edilmesi

Kaynak: (Geitgey, 2017).

Yalnızca paralel çizgileri koruyan döndürme ve ölçekleme gibi temel görüntü dönüşümleri ile Şekil 2.15'te gösterildiği gibi görüntü ortalanmakta, döndürülmekte,

ölçeklendirilmekte ve kesilerek ideal yüz görüntüsü elde edilmektedir (Geitgey, 2017).

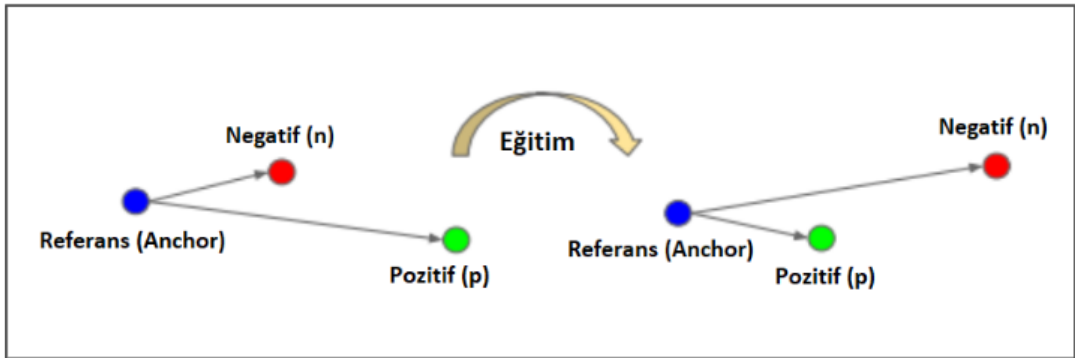


Şekil 2.15. Yüz hizalama ve ideal yüz görüntüsü elde etme

Kaynak: (Geitgey, 2017).

2.2.3 Yüze Ait Özellik Vektörü Elde Etme

Derin sinir ağına verilmeden, yani eğitim aşamasından önce her yüz için sayısal olarak ifade edilebilecek vektörler elde edilmektedir. Bu aşamada Şekil 2.16'da temsili olarak gösterilen üçlü kayıp (triple loss) yöntemi oldukça yaygın olarak kullanılmaktadır. Google araştırmacıları tarafından geliştirilen derin yüz tanıma (FaceNet) (Schroff vd., 2015) modeli bu yöntem ile kendinden önceki tüm modellerden daha başarılı sonuçlar vermiştir (Yıldırım, 2019).



Şekil 2.16. Üçlü kayıp modeli

Kaynak: (Schroff vd., 2015).

Bu yöntemde kullanılan algoritma, kendisiyle aynı olan bir yüz görüntüsünü yakınlaştırırken kendisinden farklı olan yüz görüntüsünden uzaklaştıracak şekilde sinir ağını ayarlamaktadır (Schroff vd., 2015).

Üçlü kayıp modeli, sistemin eğitilmesi sırasında test edilecek bir kişiye ait yüz görüntüsünü temel görüntü olarak kabul ederek çalışmaya başlamaktadır. Buna

“anchor” adı verilir. Aynı kişi ile etiketlenmiş başka bir görüntü daha sisteme dâhil edilerek karşılaştırma yapılmaktadır. Kendi sınıfı ile aynı sınıftan olan görüntüye pozitif görüntü adını verilir. Daha sonra temel görüntü olarak alınan yüzden farklı bir görüntü sisteme dâhil edilerek negatif görüntü olarak adlandırılır. Referans görüntü ile pozitif görüntü derin sinir ağından geçirilir ve oluşan 128 boyutunda gömme (embedding) vektörleri karşılaştırılır. Aynı şekilde negatif görüntü de derin sinir ağından geçirilir ve referans görüntü ile arasındaki uzaklık hesaplanır. Eğer referans görüntü ile negatif görüntü arasındaki uzaklık, pozitif görüntü ile arasındaki uzaklıktan küçükse ağ tam olarak eğitilememiş demektir. Geri yayılım algoritması ile parametreler güncellenerek bu işlem negatif görüntü uzaklığı pozitif görüntüden büyük oluncaya kadar devam ettirilmektedir (Yıldırım, 2019). Sonuçta sinir ağı, her bir yüz görüntüsü için 128 boyutlu bir özellik vektörü elde etmektedir.

Her yüz için sinir ağından elde edilen özellik vektör boyutunun sadece 128 olarak ele alınması gerekmemektedir. Bu konu ile ilgili meşhur FaceNet makalesinde daha farklı boyutlarda elde edilen vektörlerin başarımlarını karşılaştırması yapılmış ve 128 boyutlu vektör uzunluğu ile en yüksek başarıma ulaşıldığı belirtilmiştir (Schroff vd., 2015). FaceNet modelinde çalışılan farklı vektör boyutlarına göre başarımlar oranları Tablo 2.1’de sunulmuştur.

Tablo 2.1. Boyutları farklı özellik vektörlerinin başarımlarını karşılaştırması

Boyut	Başarımlar (%)
64	86.8
128	87.9
256	87.7
512	85.6

Kaynak: (Schroff vd., 2015).

Daha büyük vektör boyutlarının da en az küçük olanlar kadar iyi performans göstermesi beklenebilir, ancak aynı doğruluğu elde etmek için daha fazla eğitim yapmak gerekebilecektir. Küçük bir doğruluk kaybıyla daha küçük vektörlerle çalışmak da mümkündür ve **mobil cihazlarda** kullanmak için bu küçük vektör boyutları tercih edilebilir (Schroff vd., 2015).

2.3 Yüz Tanıma Sistemleri Üzerine Yapılan Araştırmalar

Literatürde yüz tanıma sistemlerinin doğruluğu ve performansı hakkında birçok araştırma yapılmıştır. 1997'de Lawrence ve ark. (Lawrence vd., 1997) ilk olarak yüzleri tanımak için CNN kullanan bir AI tekniği sunmuştur. 2010'lardan bu yana, DL tabanlı sistemlerin performansında ve doğruluğunda çok önemli kazanımlar elde edilmiştir. 2014 yılında Yaniv, T. ve ark. (Yaniv vd., 2014) Facebook araştırma laboratuvarında geliştirilen DeepFace adlı yüz tanıma sistemini sunmuştur. Bu sistem, yüz tanıma için derin mimari esasına dayanmaktadır. Bu mimaride 120 milyondan fazla bağlantılı ağırlığa sahip dokuz nöral ağ katmanı kullanılmıştır. CNN ağını internetten oluşturulan 4,4 milyon etiketli görüntüden (4000 farklı kişi) oluşan o zamana kadar mevcut olan en büyük veri kümesi ile eğitmişlerdir. Bu ağ mimarisi, DeepFace'i oluşturmak için kullanılmıştır. Bu ağ, Öklid mesafesi kullanarak karşılaştırılan tanımlayıcıları elde etmek için aynı CNN ağını bir çift yüze uygulayarak çalışmaktadır. Eğitimin amacı, aynı yüz çifti arasındaki Öklid mesafesini en aza indirmek ve 2 farklı yüzün mesafesini en üst düzeye çıkarmaktır, buna metrik öğrenme adı verilmektedir (Sufiyan, 2018; Ananthakumar, 2018; Vu vd., 2021).

2015 yılında, Google'daki AI araştırma grubu, çok büyük ölçekte uygulanabilen bir yüz tanıma sistemi olan FaceNet'i sunmuştur (Schroff vd., 2015). Bu model DeepFace'e çok benzemektedir. Ancak FaceNet, aynı kişinin iki yüzü arasındaki Öklid mesafesini en aza indiren ve farklı bir kişiye ait üçüncü bir yüz ile ilk iki yüz arasındaki Öklid mesafesini en üst düzeye çıkaran "üçlü tabanlı" (triplet-based) bir kayıp fonksiyonu kullanmaktadır. Bu yöntem, tüm son teknoloji yöntemler arasında LFW (Huang, 2008) veri kümesinde %99,63 doğruluk oranı ile en iyi performansı elde etmiştir (Sufiyan, 2018).

2016 yılında B. Amos ve ark. (Amos vd., 2016) tarafından yüz tanıma kitaplığı adı verilen OpenFace modeli tasarlanmıştır. Bu model, Google'ın FaceNet mimarisini ve önceden eğitilmiş açık kaynak kitaplıklarını birleştirmektedir. Sistem, LFW veri kümesi üzerinde %92,92 doğruluk oranıyla test edilmiştir. Daha küçük eğitim veri setine rağmen, bu model LFW doğrulama karşılaştırmasında rekabetçi doğruluk ve performans sonuçları göstermektedir.

2016 yılında Davis King (King, 2017) tarafından oluşturulan ve 2017 yılında Adam Geitgey (Geitgey, 2017) tarafından geliştirilen açık kaynaklı bir yüz tanıma

kütüphanesi olan Dlib modeli piyasaya sürülmüştür. Bu model, birkaç katmanın kaldırıldığı ve katman başına filtre sayısının yarı yarıya azaltıldığı 29 evrişim katmanına sahip bir ResNet-34 ağının bir sürümünü kullanmaktadır. Bu kitaplık, 3 milyon görüntü ile eğitilmiş son teknoloji yüz tanıma sistemiyle oluşturulmuştur. LFW veri kümesinde %99,38 doğruluk oranına sahiptir. Her yüz görüntüsü için 128 boyutlu bir vektör uzayıyla temsil edilebilen gömmeyi (embedding) kullanmaktadır.

2022'de Chu ve arkadaşları (Chu vd., 2022) tarafından yayınlanan sahne değişiklikleri modülü yardımıyla derin metrik öğrenme tabanlı bir yüz tanıma sistemi geliştirilmiştir. Çalışmalarında, yüz tanıma hedefine ulaşmak için video görüntülerindeki sahne değişim durumlarını tespit eden ve yüz için yeniden tanınması gerekmeyen kareleri atlayan CFSM modelini önermişlerdir. Modele göre gömülü sistemdeki gerçek zamanlı yüz tanıma sisteminin iş yükünü önemli ölçüde azaltacaklarını belirtmişlerdir. Uygulama, yüz tanıma ve yüz çıkarma aşamaları için MTCNN ve FaceNet sinir ağı modellerini kullanan bir mekanizmaya sahiptir. Modeli Youtube-8M veri setinden 1280×720 piksel kare boyutlarına sahip on adet video klip ile test etmişler ve ortalama 0,896 doğruluğa ulaşmışlardır (Chu vd., 2022).

Derin metrik öğrenmeye dayalı genel kabul görmüş bazı çalışmalar ve öne çıkan özellikleri aşağıdaki tabloda özet olarak verilmiştir.

Tablo 2.2. LFW veri seti üzerinde uygulanan bazı mimarilerin karşılaştırması

Mimari	Yıl	Başarım (%)	Öne Çıkan Özellikler
DeepFace	2014	97.35	Geniş bir veri seti ile oluşturulan CNN mimarisini kullanarak ilk kez derin metrik öğrenme kavramını sunar.
FaceNet	2015	99.63	Son teknoloji yöntemler arasında en iyi performansı vermiştir.
OpenFace	2016	92.92	Daha küçük eğitim veri setine rağmen makul doğruluk ve performans sonuçları verir.
Dlib	2017	99.38	Derin ağ modelinden filtre ve katman sayısı azaltılarak hesaplama maliyeti de azaltılmıştır.

Kaynak: (Serengil, 2021).

Tablo 2.3.'te ulusal tez merkezinden elde edilen ülkemizdeki yüz tanıma konusu ile ilgili yazılmış bazı doktora tez çalışmaları özet olarak gösterilmektedir.

Tablo 2.3. Ülkemizde yazılan yüz tanıma konusu ile ilgili bazı doktora tezleri

Yazar	Yıl	Tez Adı
MURAT TAŞKIRAN	2022	Yüz ifadelerinin görünüm tabanlı ve dinamik özellikleri kullanılarak olumsuz koşullar altında hibrit yüz tanıma
ÖMER FARUK SÖYLEMEZ	2021	Yüz ifadesi tanıma başarımı için ilgili bölge belirleme ve çekişmeli üretici ağlar kullanılarak örnek artırım yöntemlerinin geliştirilmesi
EMRAH BAŞARAN	2020	Kişi tanıma için yüz tanıma ve kişinin yeniden tanınması
FATİH ŞAHİN	2020	Simgesel noktaları kullanarak şekil ve gri düzeyde görüntüleme dayalı yüksek ayırmacı yüz tanıma yöntemi
FARUK AYATA	2020	İçerik tabanlı görüntü erişim yöntemleriyle aile bireylerinde yüz tanıma sistemi
SAAD OMRAN ELHASHMI ALLAGWAIL	2019	Orijinal ve simetrik örnekleri kullanarak yüz tanıma için yeni ve etkili bir yöntem
İSMAİL ÖZTEL	2018	Kısmi ve tam yüz görüntüleri üzerinde makine öğrenmesi yöntemleriyle yüz ifadesi tespiti
MERYEM UZUN PER	2018	Yerel Walsh dönüşümü ile yüz tanıma
ABUBAKAR MUHAMMAD ASHIR	2018	Uyarlanabilir sınıflandırma ile yüz ifade tanıma için yeni bir yaklaşım
NEŞE ALYÜZ	2013	Örtme durumunda üç boyutlu yüz tanıma
KRİSTİN SURPUHİ BENLİ	2013	Yüz anatomisine dayalı ifade tanıma
GÜLSÜM ÇİĞDEM ÇAVDAROĞLU	2013	Sanal gerçeklikte ilinti operatörleri geliştirilerek yüz tanıma analizi
NESLİ ERDOĞMUŞ	2012	Utilization of 3D data in face recognition
BİRKAN TUNÇ	2012	Doğrusal olmayan manifoldlar üzerinde gürbüz yüz tanıma
TOLGA İNAN	2011	Yerel şekil betimleyiciler ile üç boyutlu yüz tanıma
GÖKSEL GÜNLÜ	2010	3 boyutlu ayırık kosinüs dönüşümü tabanlı yüz bulma ve tanıma
ALAA ADNAN ELEYAN	2009	Face recognition from still images and video sequences
FAHRİ TUNÇER	2008	Küresel harmoniklerle 3 boyutlu yüz modelleme ve tanıma
ÜMİT ÇİĞDEM TURHAL	2008	İki boyutlu yüz tanıma metodlarına yeni yaklaşımlar; satır ve sütun vektörleri arasındaki değişimlerin kullanılması
ALBERT ALİ SALAH	2007	Biyolojik tabanlı üç boyutlu yüz tanıma
BERK GÖKBERK	2006	Üç boyutlu yüz tanıma
GÜLESER KALAYCI	2001	Yüz tanıma uygulamalı doğrusal ayırtaç analizi için çevrimiçi yerel algoritmalar

Kaynak: <https://tez.yok.gov.tr/UlusalTezMerkezi/>

Bu tez çalışmasında, kayan pencerelerde meydana gelebilecek hatalı tanımanın azaltılması veya yok edilmesi için önerilen sahne değişim göstergesinden faydalanarak ortaya yeni bir model ve bakış açısı getirilmeye çalışılmıştır. Bu model, yüz tanıma aşamasında kullanılan yakınlık değerlerini dinamik olarak değiştirerek hatalı tanımlama problemlerinin azaltılmasına fayda sağlamak üzere tasarlanmıştır. Literatürdeki mevcut modellerin başarımları oranları incelendiğinde birbirine oldukça yakın ve yüksek başarımlara ulaşıldığı görülmektedir. Bu yüzden, mevcut

modellerden Dlib modeli tercih edilerek gerçek zamanlı yüz tanıma sistemi bu model üzerinde geliştirilmiştir. Bu çalışma ile daha önce Güllü ve ark.'larının (Güllü vd., 2004) yapmış olduğu sahne geçişlerini algılama çalışmasından ilham alınarak, önerilen yüz tanıma sisteminde sahne değişikliklerinin ayrı bir parametre olarak kullanılmasının yüz tanıma meydana gelebilecek hataların azaltılmasına katkı sağlayacağı savunulmaktadır.

2.4 Yüz Tanıma Sistemlerinde Karşılaşılan Bazı Problemler

Yüz tanıma sistemlerinin başarımı sistemin bir bütün olarak en verimli şekilde çalışmasına bağlıdır. Sürecin her aşamasında başarımı olumsuz etkileyecek durumlar meydana gelebilmektedir. Bunlardan ilki yüz tespiti aşamasıdır. Bu aşamada yüz sınırları doğru bir şekilde tespit edilemezse sistem hatalı davranışlar sergileyebilmektedir. Yüz tespitinin yapılamadığı veya hatalı tespite yol açabilecek bazı durumlar genel olarak aşağıdaki ana maddeler halinde özetlenebilir.

- Poz açısı kaynaklı durumlar
- Yüz tespit algoritması kaynaklı durumlar
- Yüz boyutu ile ilgili durumlar
- Ten rengi gibi etmenler
- Yüz ifadelerinden veya yüzdeki değişikliklerden kaynaklı durumlar
- Işık yönü veya şiddeti ile ilgili durumlar

Aşağıdaki Şekil 2.17(a)'da yüz ana hatları açık bir biçimde görünmesine rağmen, tercih edilen algoritma tarafından yüz tespiti yapılamamıştır. Meydana gelen bu problem yüz boyutunun küçük olmasından kaynaklı olabileceği gibi yüz dışındaki ortam farklılıklarından kaynaklı da olabilmektedir. Şekil 2.17(b) ve Şekil 2.17(c)'de yüzün tanınmasını büyük ölçüde engelleyen nesnelerin probleme sebep olduğu söylenebilir. Çünkü Şekil 2.17(d)'de bu nesnelerin bulunmadığı görüntülerdeki tüm yüzlerin tespit edilebildiği ve böylece başarılı bir biçimde yüz tanıma işlemi gerçekleştirildiği gözükmektedir.



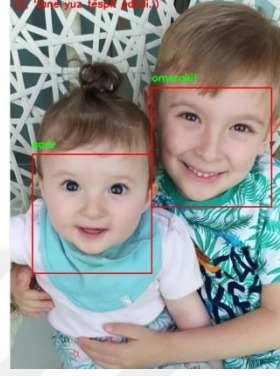
a)



b)



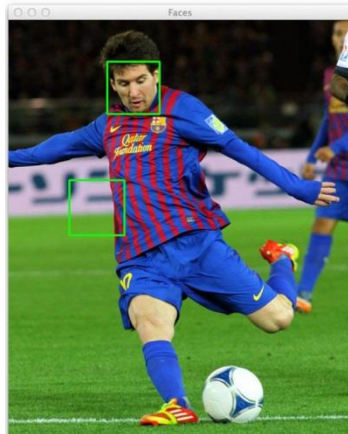
c)



d)

Şekil 2.17. Yüz tespitini zorlaştıran örnek durumlar a) Küçük boyutlu yüz görüntüsü b) ve c) Yüzde nesne olması durumu d) İdeal yüz tespiti.

Şekil 2.18(a) ve (b)'de ise yüz olan bölgede yüz tespit edildiği gibi, yüz içermeyen bölgede de yüz varmış gibi hatalı sonuç üretilmiştir.



a)



b)

Şekil 2.18. Yüz bulunmayan bölgede yüz tespiti durumu

Şekil 2.19(a) ve (b)'de ise yüzün büyük bir kısmını kapatan nesne olmasından dolayı yüz tespit algoritması yeterli sayıda yüz işaret noktası elde edememiş ve yüz tespiti yapılamamıştır.



a)

b)

Şekil 2.19. Yüz işaret noktalarının elde edilememesi durumu

Şekil 2.20’de ise çok fazla ve uygun açıda olmayan yüz görüntüsü içeren kalabalık bir ortamda birden fazla yüzün tespit edilemediği örnek bir durum gözükmektedir.



Şekil 2.20. Uygun açıda olmayan kalabalık ortam yüz görüntüleri.

Bu hatalı durumlar sistemin eğitilmesi aşamasında ortaya çıktıysa, eğitim hataları veya aykırı veri olarak ele alınarak ayıklanabilir. Gerçek zamanlı olarak sistemin test edilmesi esnasında meydana gelmesi durumunda, boyut üzerinde değişiklikler yapılabilir, yüz tespit algoritmasının değerleri ile hassasiyet oranları düzenlenebilir veya örneğin CNN tabanlı yüz tespit algoritması gibi başka bir tespit metoduna başvurulabilir. Sadece yüz tespiti konusu ile ilgili Literatürde başlı başına birçok çalışma mevcuttur. Bu tez konusu kapsamında yüz tespiti alanının inceliklerine değinilmeden genel hatları ile oluşan belli başlı problemlerden bahsedilmiştir.

ÜÇÜNCÜ BÖLÜM

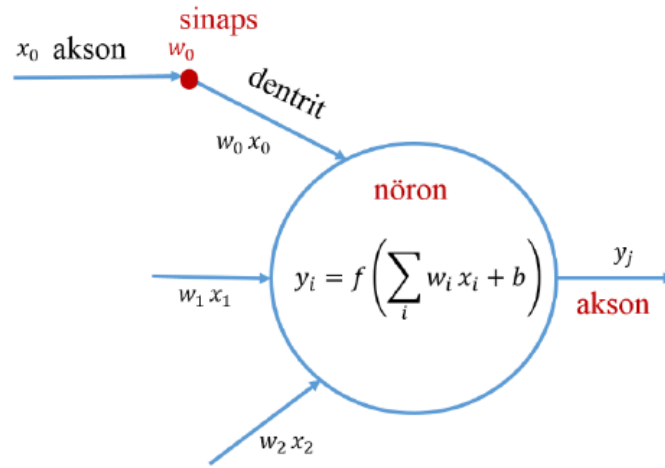
DERİN ÖĞRENME

3.1 Derin Öğrenme Nedir?

Hinton tarafından ortaya atılan derin öğrenme kavramı, yapay sinir ağlarına yeni bir bakış açısı getirmiştir ve derin konvolüsyonel (evrişimsel) sinir ağları ismini buradan almaktadır (Hinton vd., 2006).

Derin evrişimsel sinir ağları aslında çok fazla katmandan oluşan sinir ağları olarak ifade edilebilir. Derin evrişimli sinir ağları ile diğerlerinden çok daha yüksek başarılı sonuçlar elde edilmiştir (Krizhevsky vd., 2012; LeCun vd., 1998; Doğan ve Türkoğlu, 2019).

Temel bir yapay sinir ağı yapısı aşağıdaki şekilde gösterilmektedir.



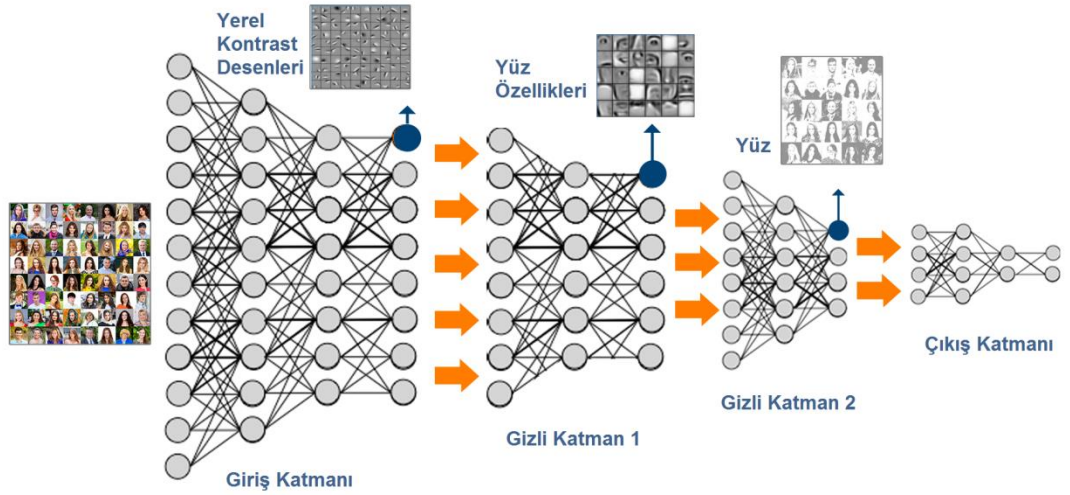
Şekil 3.1. Yapay sinir ağı yapısı

Kaynak: (Hanilçi, 2019).

Yapay sinir ağlarının ve olasılıksal grafik modellerinin simgesi olan DÖ yöntemleri, giriş verilerine ait doğrusal öznitelikleri çıkarmak yerine gizli katmanlarda daha karmaşık ilişkilerin elde edilebilmesine olanak sağlamaktadırlar. Derin öğrenme yaklaşımı, nesne tanıma, konuşma ve doğal imgelerin incelenmesi gibi birçok alanda başarıyla kullanılmıştır. Özellikle ilk başlarda MNIST el yazısı karakter tanıma alanında ve daha sonra ImageNet veri kümelerinde elde edilen yüksek sınıflandırma

başarımı ile ön plana çıkmıştır. Derin öğrenmede Kısıtlı Boltzmann Makineleri ve Otomatik Kodlayıcılar başarılı bir şekilde kullanılan yöntemlerdir. Bu yöntemlerin insan yüzlerindeki istatistiksel yapının öğrenilmesine imkân sağlayacağı öngörülmektedir (Özcan, 2014).

DÖ ile veri üzerinde ön işlem ve elle özellik çıkarımı gibi işlemlere gerek olmadan birçok işlem bir arada yürütülmektedir. Bu işlemler DÖ içindeki gizli katmanlar yardımıyla otomatik olarak gerçekleştirilmektedir. Bu katmanlar hiyerarşik bir yapı oluşturarak ayrıca özellik çıkarmak için başka bir alana ihtiyaç duymamaktadır. Belirgin özellikler kademeli olarak tespit edilmekte ve sonraki katmanlara iletilmektedir (Hinton ve Salakhutdinov, 2006; Bengio, 2009). Şekil 3.2’de temsili bir derin evrişimli sinir ağı yapısı gösterilmektedir.



Şekil 3.2. Derin sinir ağı yapısı

Kaynak: (Sharma, 2018)

3.2 Evrişimli Sinir Ağları

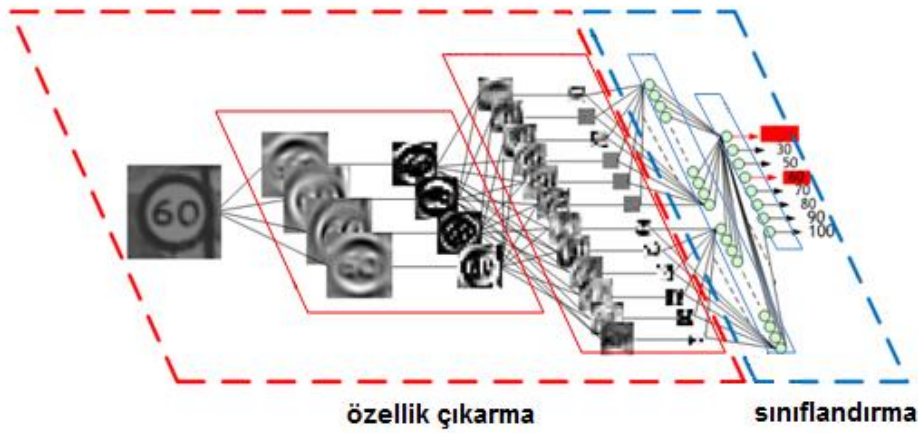
İlk olarak 1998 yılında LeCun ve arkadaşları tarafından geliştirilerek gradyan temelli bir yaklaşım ile ortaya çıkarılmıştır (LeCun vd., 1998).

Evrişimli Sinir Ağları bir çeşit çok katmanlı algılayıcılar olarak düşünülebilir. Görüntülerdeki kenar gibi basit hücre özelliklerinin oluşması için görsel alt bölgelere, karmaşık hücreler için ise daha derin alıcılarla yoğunlaşmaktadır. ESA algoritması, ileri yönlü bir sinir ağı olarak ifade edilebilir. Evrişim işlemi ile

gerçekleşen matematiksel işlem, bir nöronun kendini uyarılara verdiği cevap olarak ifade edilebilir (Hubel ve Wiesel, 1968; Fukushima, 1980).

ESA, evrimsel katmanlar, alt örnekleme (subsampling) katmanları ve son aşamada tamamen bağlı (fully connected) katmanlardan oluşmaktadır (Song ve Lee, 2013).

ESA'da evrişim katmanı, nonlineerlik katmanı ve seyreltme (pooling) katmanları ile basit temel özneliklerden karmaşık özneliklere doğru hareket etmektedir. Her katman kendi işlevini yürütmekte ve sınıflandırıcı katmanında ise sonuç üretilmektedir. ESA'da evrişim katmanlarında kullanılan filtreler eğitim boyunca hatayı minimize edecek şekilde kendilerini güncellemektedirler. Filtreler güncellendikçe katman çıktılarından elde edilen özneliklerde farklılıklar oluşmaktadır. Eğitim aşaması boyunca bu tekrarlı iyileşme daha ayırt edici özneliklerin çıkarımını sağlamaktadır (Bilgiç vd., 2017). Özellikle büyük verilerin işlendiği durumlarda bu yöntemin başarısı kanıtlanmıştır (Çalık vd., 2017; Ranzato vd., 2007; Chen ve Lin, 2014). Şekil.3.3'de temsili olarak bir ESA mimarisi gösterilmiştir.



Şekil 3.3. ESA mimarisi

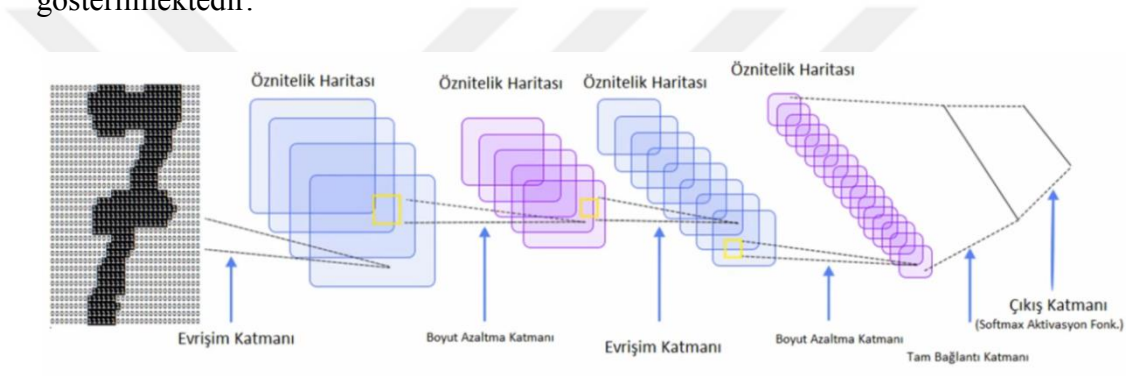
Kaynak: (Peemen, 2018).

Sınıflandırma, doğrulama gibi işlemler için ESA sıklıkla kullanılmaktadır. Model, iki ana aşama olmak üzere, içerisindeki katmanlardan elde edilen temel özneliklerin ortaya çıkartılması ve sınıflandırma yapılması şeklinde özetlenebilir. Konvolüsyonel sinir ağları, 2012 ve 2014 yılında yapılan Imagenet (Büyük Ölçekli Görsel Tanıma)

yarışmasında önemli başarılar elde ettiği için oldukça popüler hale gelmiştir (Doğan ve Türkoğlu, 2019; Goodfellow vd., 2013).

ESA'nda aynı katman seviyelerinde bulunan birimler aynı ağırlıkları paylaşacak şekilde sıralı olarak gruplandırılmaktadır. Her katman grubunun çıktı kümesine özellik haritası (feature map) adı verilmektedir. Bu oluşan özellik haritasındaki gruplar, aynı filtreleme işlemi ile görüntü üzerinde sürekli gezinerek farklı bölgelere rasgele uygulanmaktadır. Verimli bir evrişim katmanı ile birden fazla özellik haritası elde edilmekte ve böylece her katmanda birbirinden farklı özelliklerin çıkarılması sağlanmaktadır (Şahin, 2017).

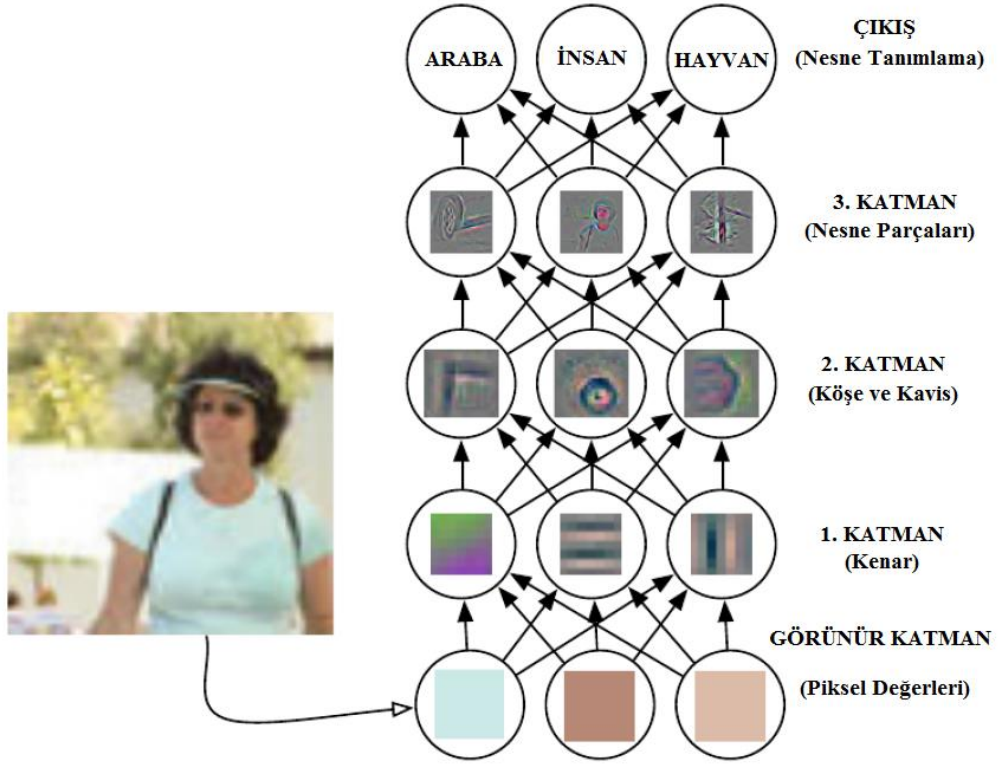
Şekil 3.4'te temel ESA katmanları ve oluşan öznelik haritaları temsili olarak gösterilmektedir.



Şekil 3.4. ESA temel katmanları

3.3 Derin Öğrenme Ağları İşlemler ve Katmanları

Derin öğrenme ile bir görüntünün tanımlanmasında kullanılabilecek basit temsili özelliklerden olan kenarlar, köşeler ve kavisler gibi nesnelere oluşturan parçaları içerip içermediğine bakılarak özellik haritaları elde edilmektedir. Şekil 3.4'teki gibi katmandan katmana bağlanarak temel şekilsel özellikler birbirleri ile bağlantılı olarak elde edilmektedir (Badem, 2017). Şekil 3.5 ile katmanlar arasında oluşan özelliklere ait temsili bir gösterim sunulmuştur.



Şekil 3.5. ESA temel özelliklerin ortaya çıkarılması

Kaynak: (Badem, 2017).

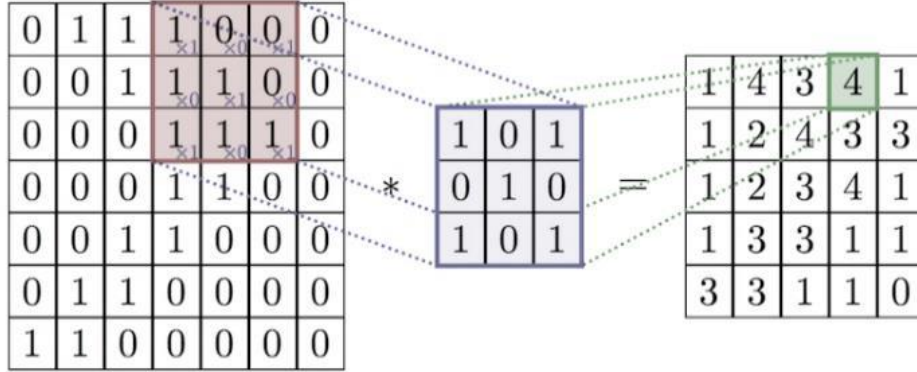
3.3.1 Giriş Katmanı

Bu katmanda veriler ham olarak modele girdi olarak verilmektedir. İşlenecek verinin boyutu modelin başarımı için önemli bir parametredir. Yüksek boyutlu veriler, daha çok bellek kullanım ihtiyacını ortaya çıkararak eğitim sürelerini olumsuz şekilde etkilemektedir. Buna karşın ağ başarımına ise olumlu katkılar sunabilecektir. Boyutun düşük seçilmesi bellek ihtiyacını azaltarak eğitim süresini kısaltmaktadır. Ancak, buna rağmen modeldeki ağın derinliği azalırken başarımları daha düşük seyredebilmektedir (İnik ve Ülker, 2017).

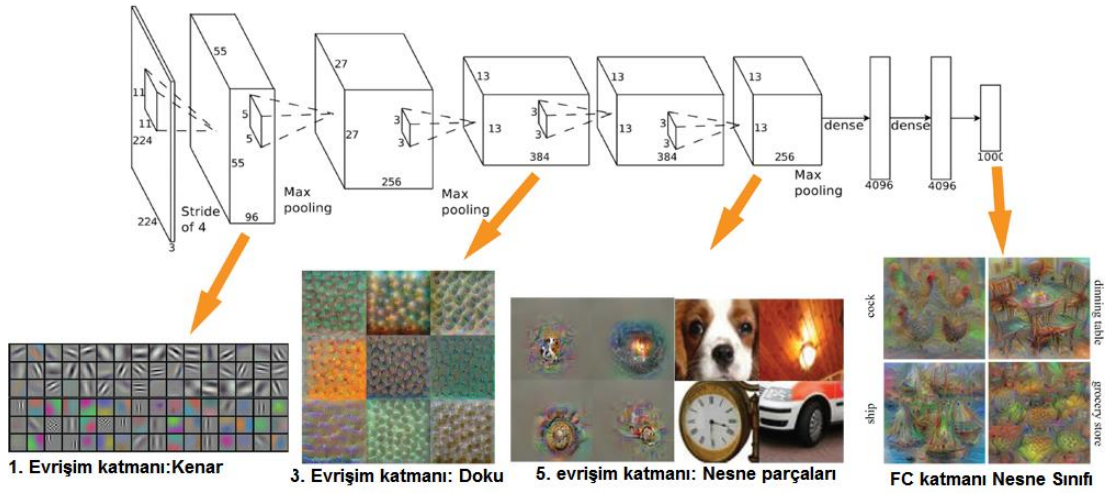
3.3.2 Evrişim (Konvolüsyon) Katmanı

Bu katman, ESA yapısının temelini oluşturan filtreler ile dönüşüm işlemi sağlamaktadır. Bu filtreler tüm görüntü üzerinde hareket ederek farklı özelliklerin ortaya çıkarılmasında etkili olmaktadır. 2x2, 3x3, 5x5 gibi farklı boyutlardan oluşan filtreler rasgele değişerek katmanlı mimariyi meydana getirmektedir. Bu filtreler ile önceki katmandan gelen görüntülere evrişim işlemi uygulanmakta ve çıkış verisi elde

edilmektedir. Şekil 3.6’da evrişim işlemine ait bir örnek gösterilmektedir. Bu evrişim süreci sonunda aktivasyon haritası (özellik haritası) oluşmaktadır (İnik ve Ülker, 2017). Şekil 3.7’de her katman sonrası ortaya çıkarılan özelliklere ait bir temsili gösterim sunulmuştur.



Şekil 3.6. Evrişim işlemi

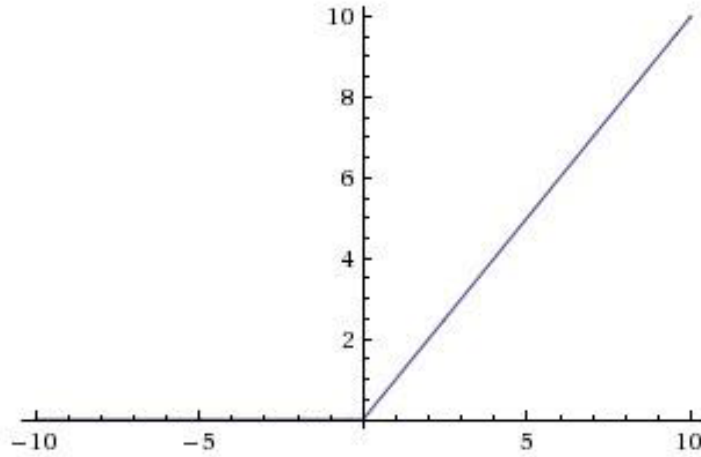


Şekil 3.7. ESA katmanlarından elde edilen özellikler

3.3.3 Aktivasyon Katmanı

Genellikle evrişim katmanından sonra aktivasyon katmanı gelmektedir. Aktivasyon fonksiyonu olarak çeşitli fonksiyonlar kullanılmaktadır. Bunlardan bazıları, hiperbolik tanjant, ReLU, sigmoid, sinüs, adım (step) ve eşik değer fonksiyonlarıdır. Bu fonksiyonlardan ReLU oldukça yaygın kullanım alanına sahiptir (Doğan, F. ve Türkoğlu, İ., 2018).

Aşağıda ReLU aktivasyon fonksiyonuna ait bir grafik gösterilmektedir.



Şekil 3.8. ReLU aktivasyon fonksiyonu

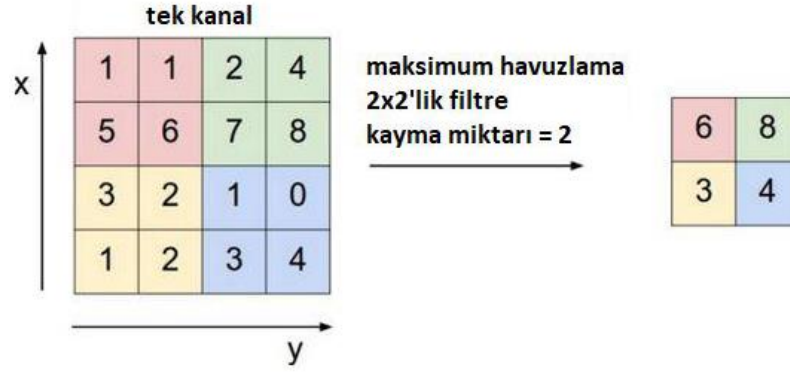
Kaynak: (Öznel, 2019).

ReLU ile aktivasyon işlemine giren giriş değeri negatif ise 0 değerini, pozitif ise giriş değerini döndürmektedir. Bu katmanın kullanılması ile modelin daha hızlı eğitilmesi sağlanmaktadır.

3.3.4 Havuzlama (Pooling) Katmanı

Havuzlama katmanı da diğer katmanlar gibi evrişimsel ağın önemli bir parçasıdır. Bu katmanın temel görevi evrişim katmanlarından çıkan verinin boyutunu azaltmak ve bu sayede hesaplama maliyetini düşürmektir. Farklı özellik haritalarına ayrı ayrı havuzlama işlemi uygulanabilmektedir. En çok maksimum ve ortalama havuzlama yöntemleri tercih edilmektedir. Bu yöntemlerle daha küçük matris boyutları elde edebilmek için iki boyutlu giriş verisi eşit boyutta matrislere ayrılarak oluşturulan alt matrislerin değerleri seçilen yönteme göre dönüşüme uğratılmaktadır (Şahin, 2017).

Şekil 3.9'da maksimum havuzlama yöntemi ile dönüştürülen örnek bir havuzlama işlemi gösterilmektedir.



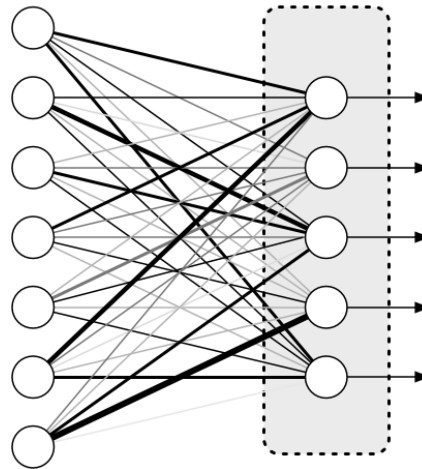
Şekil 3.9. Maksimum havuzlama örneği

Kaynak: (Öznel, 2019).

3.3.5 Tam-Bağlı (Full-Connected) Katman

Bu katman önceki katmandaki tüm nöronlara bir sıra gibi tam olarak bağlı görünmektedir. Tanımlanması istenen nesneyi belirleyecek olan özelliklerin hangi sınıfa ait olduğu bu katmanda belirlenmektedir. Aktivasyon haritalarında yüksek değere sahip olan nöronlar sınıfın belirlenmesinde ön plana çıkmaktadır. Bu yüzden bu katman belirlenecek sınıfla ilişkisi en yüksek katmandır (Doğan, F. ve Türkoğlu, İ., 2018).

Nöronlar içerisinde bu özellikleri içeren ağırlıklara bakılarak sınıf ataması yapılmaktadır (LeCun ve Bengio, 1995). Kabaca bu katman, klasik bir yapay sinir ağı katmanıdır. Şekil 3.10 ile temsili bir tam-bağlı katman yapısı gösterilmektedir.



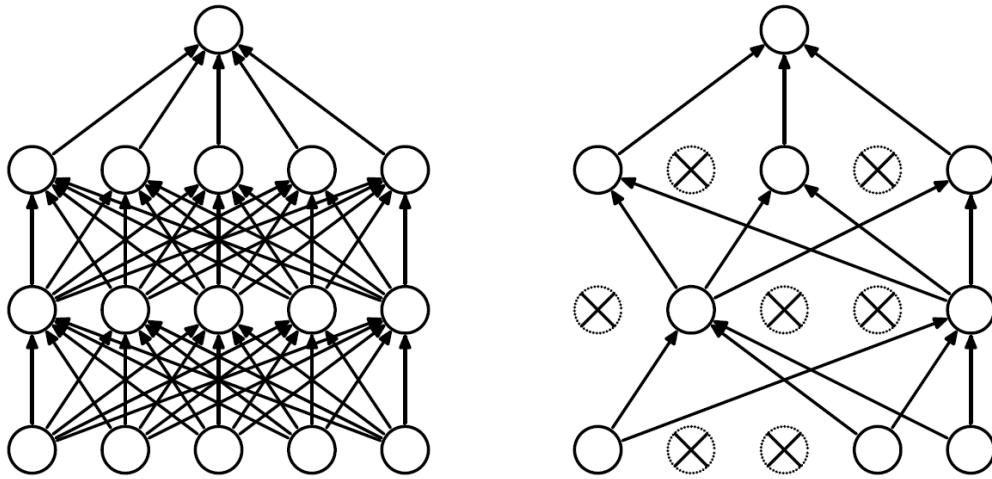
Şekil 3.10. Tam-bağlı katman yapısı

Kaynak: (Doğan, F. ve Türkoğlu, İ., 2019).

Evrişim ve havuzlama katmanlarından çıkan görüntüler tam bağlı katmanlara girdi olarak verilmektedir. Tam bağlı katmanın öğrenebilmesi için giriş verilerine düzleştirme (flatten) adı verilen vektöre dönüştürme işlemleri gerçekleştirilmektedir (Akın, 2019).

3.3.6 Dropout Katmanı

Eğitim aşamasında çok katmanlı yapay sinir ağları aşırı öğrenme adı verilen ağın ezberlenmesi olayı ile karşılaşmaktadırlar. Bu istenmeyen durumu önlemek için Hinton ve arkadaşları tarafından ağda ezberlemeye sebep olabilecek bazı düğümlerin ortadan kaldırılması önerilmiştir (Şahin Ö. , 2017). Bu işleme dropout adı verilmektedir. Her eğitim durumu için nöronların yarısını rastgele atlayarak çalışmaktadır. Yani, eğitim aşamasında bu işlem ile ağ boyutu azaltılmış olmaktadır. Bu işlem, nöronları birbirleri ile daha uyumlu hale getirmektedir (Srivastava vd., 2014). Aşağıdaki şekilde örnek bir dropout işlemi görsel olarak verilmiştir.



Şekil 3.11. Dropout işlemi

Kaynak: (Şahin Ö. , 2017).

3.3.7 Sınıflandırma ve Yumuşatma (Softmax) Katmanı

Bu katman, derin öğrenme modelinin son katmanı olarak, tam bağlantılı katmandan sonra gelen ve sınıflandırma işleminin yapıldığı katmandır. Bu katmanın çıkış nöron değeri, sınıflandırılacak objelerin sayısı ile ifade edilmektedir. Örnek olarak, 15 farklı nesnenin sınıflandırılması isteniyorsa, çıkış değeri 15 nöron olmalıdır. Bu

son katmanda çoğunlukla softmax aktivasyon fonksiyonu kullanılmaktadır (İnik ve Ülker, 2017).

3.3.8 Normalizasyon (Normalization) İşlemi

Derin konvolüsyonel sinir ağlarının eğitilmesi yoğun hesaplamalar gerektiren ciddi bir süreç almaktadır. Eğitim süresinin azaltılması için nöronların aktivasyonları normalize edilmelidir. Normalizasyon işlemi geri beslemeli ağlarda gizli katmanlardaki durumları kararlı hale getirmekte oldukça etkilidir. Genellikle ReLU aktivasyon fonksiyonundan sonra normalleştirme işlemleri gerçekleştirilmektedir (Li vd., 2015).

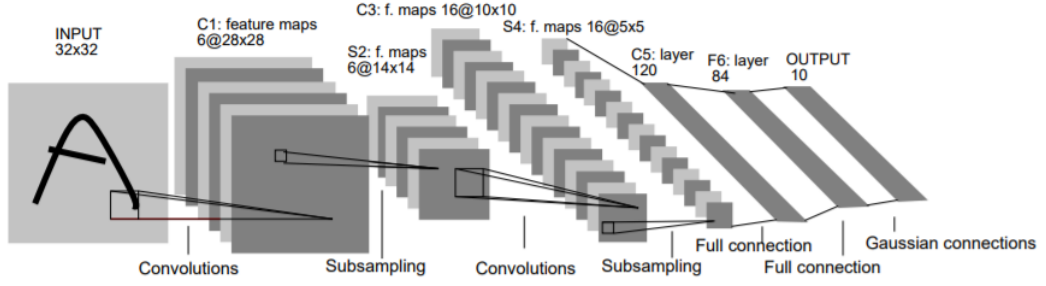
3.4 Derin Öğrenme Temel Algoritmaları

ESA'lar, güncel çalışmalar dikkate alındığında görüntü sınıflandırma alanında oldukça başarılı sonuçlar vermektedir. Görüntü sınıflandırma problemleri için klasik görüntü işleme teknikleri yerine ESA kullanmak daha verimli sonuçlar vermektedir (Şahin, 2017).

Derin sinir ağlarında kullanılmakta olan genel kabul görmüş “State of the Arts” olarak adlandırılan bazı ağ modelleri vardır. Aşağıda maddeler halinde bu ağ modellerinden ön plana çıkmış olanlar hakkında bilgiler verilmektedir. Genel olarak bu modeller, evrişim işlemi yapma biçimleri ve tercih edilen filtre boyutları, havuzlama uygulama türü, yerleri ve sayısı, daha az hesaplama yapabilmek ve dolayısıyla hafızayı daha etkin kullanabilmek için geliştirilen özel teknikler gibi farklılıklar ortaya koymuş modellerdir.

3.4.1 LeNet

İlk evrişimsel sinir ağı olarak bilinen ve Yann LeCun tarafından geliştirilen modeldir (LeCun vd., 1998). Bu model iki adet evrişim, iki adet havuzlama ve iki adet tam bağlı katmandan meydana gelmektedir. Aşağıda LeNet mimarisini ifade eden örnek bir şekil gösterilmektedir. İlk olarak MNIST veri kümesinde uygulanmıştır. Evrişim katmanlarında 5x5 boyutlu filtreler kullanarak havuzlama katmanlarında 2x2 boyutlu filtreler ile boyut azaltma yapılmaktadır. Derin sinir ağı kısmında 120 ve 84 düğümlü katmanlar kullanarak, sınıflandırma için çıkış katmanı 10 elemandan (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) oluşmaktadır (Akın, 2019).



Şekil 3.12. LeNet mimarisi

Kaynak: (Akın, 2019).

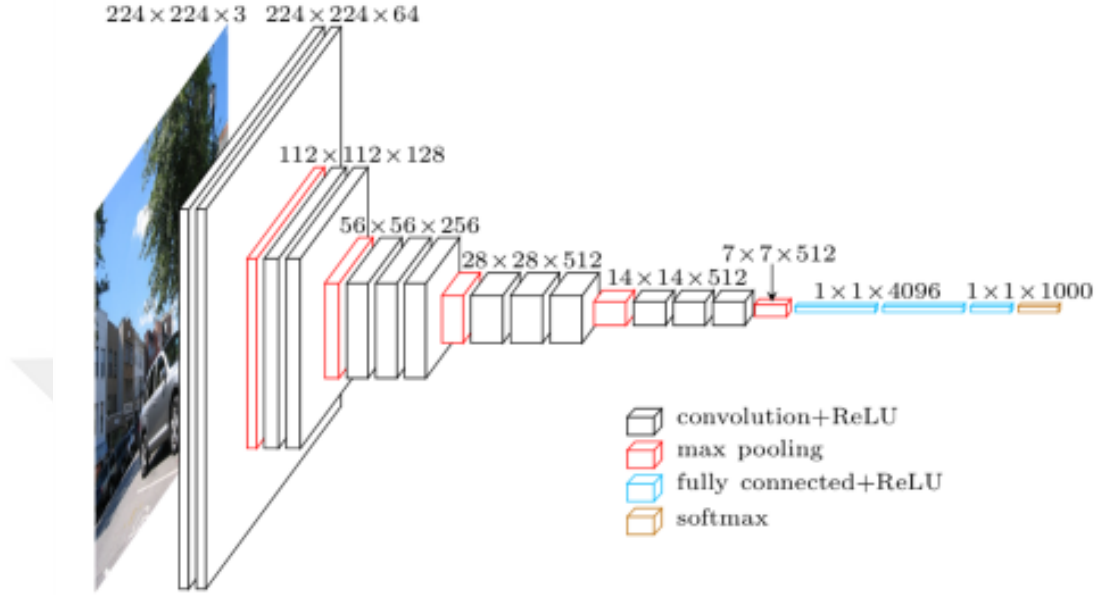
3.4.2 AlexNet

Alex Krizhevsky vd. tarafından geliştirilen AlexNet modeli 2012 yılında düzenlenen ImageNet yarışmasını kazanan modeldir (Krizhevsky vd., 2012). 5 adet konvolüsyon ve 3 adet tam bağlı katmandan oluşmaktadır. RGB formatındaki renkli bir görüntü 224x224 boyutundan kırmızı, yeşil ve mavi renk değerlerini de içerebilmesi için 3x224x224 boyutuna dönüştürülmektedir. Ağın çıkışındaki tam bağlantı katmanı 1000 nörondan oluşacak şekilde tasarlanmıştır. Böylece 1000 sınıflı bir ağ, softmax katmanı ile bu sınıflara ait olasılık dağılımını elde edebilecektir (Şahin, 2017; Akın, 2019).

Aktivasyon fonksiyonu olarak AlexNet, ReLU fonksiyonunu kullanmaktadır. ReLU aktivasyon fonksiyonu bilindiği üzere daha hızlı karar verebildiği için eğitim süresini önemli ölçüde kısaltmaktadır (Özgel G. , 2019). Bu modelde ağırlık gecikmeleri ve momentum değerleri için gradyan iniş modeli kullanılmıştır. 25 katmandan oluşan AlexNet mimarisi, ImageNet yarışmasında o ana kadar ölçülen mevcut %26,1'lik hata oranını %15,3'e indirerek önemli bir başarı elde etmiştir. Bu başarı AlexNet'i oldukça popüler DÖ mimarilerinden biri haline getirmiştir (Doğan ve Türkoğlu, 2019).

Şekil 3.13'te AlexNet mimarisi gösterilmektedir.

2014 yılındaki ILSVRC yarışmasında VGG16 modeli ImageNet verisi üzerinde eğitilerek %7,3 hata oranı ile oldukça başarılı bir performans sergilemiştir. Model, lokalizasyon kategorisinde birinci, sınıflandırma kategorisinde ise ikincilik derecelerini elde etmiştir (Şahin, 2017).



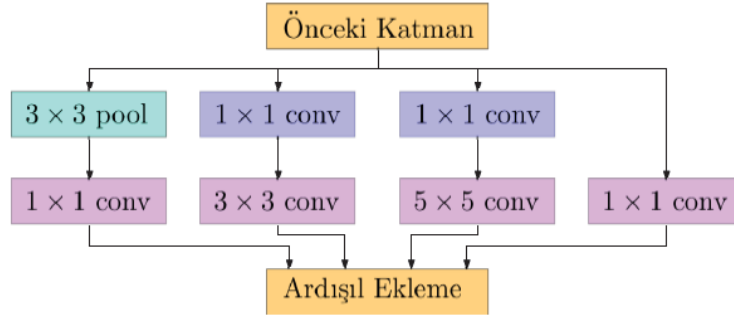
Şekil 3.14. VggNet mimarisi

Kaynak: (Rosebrock, 2018).

3.4.4 GoogleNet

GoogleNet mimarisi, Inception olarak adlandırılan bir modül yardımıyla farklı boyutlarda filtreleme yaparak yeni bir yaklaşım getirmiştir. AlexNet'in tasarımında ilk evrimsel katmanda düşük seviyeli özneteliklerin çıkartılması için 11x11 boyutlu filtreler tercih edilmiştir. Ancak giriş verisi için hangi boyuttaki filtrenin tercih edilmesinin daha verimli olabileceği bilinmemektedir. Bu fikirden ilham alarak GoogleNet ortaya sürülmüştür. Modelin karmaşık bir yapısı vardır. GoogleNet, girdi olarak adlandırılan tensöre farklı boyutlarda rasgele filtrelerin uygulanması ve çıktılarının üçüncü boyutta ele alınarak ardışıl şekilde birbirine eklenmesi ile hareket eder. Her bir evrişim katmanında ReLU aktivasyon fonksiyonu kullanılmaktadır. Giriş tensörüne ilk olarak 1x1 filtreler uygulanarak daha az derinliğe sahip alt uzaya yönlendirilmekte ve böylece ardışıl ekleme işlemi ile gelen derinliğin hızlı artışı engellenmektedir (Çalık, 2019).

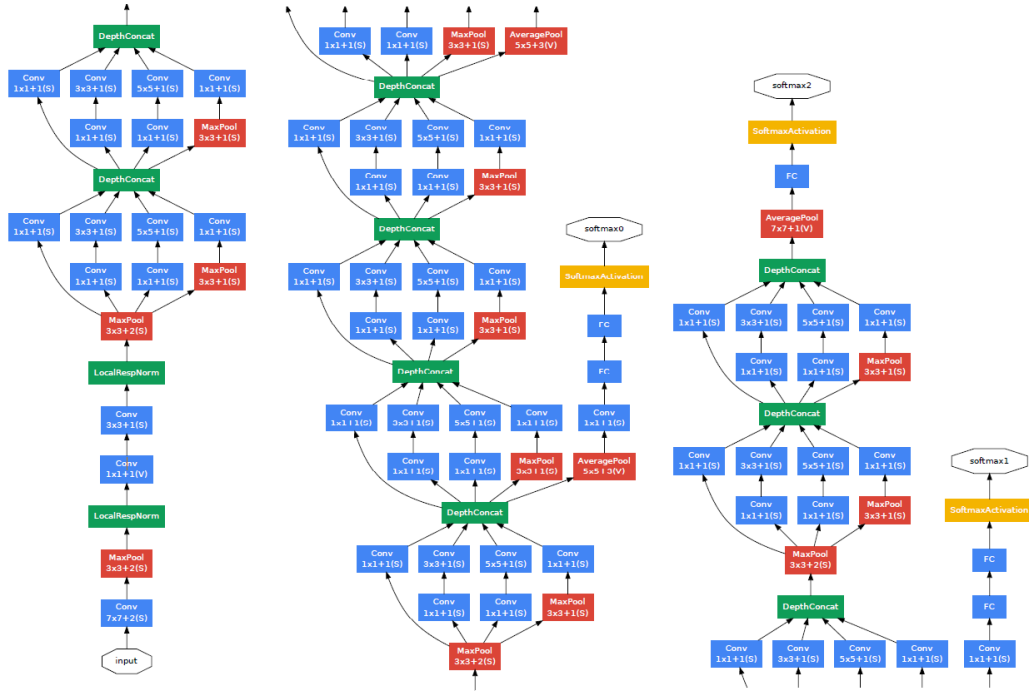
Şekil 3.15'te GoogleNet inception modülü örneği gösterilmektedir.



Şekil 3.15. GoogleNet inception modülü örneği

Kaynak: (Çalık, 2019).

ImageNet (ILSVRC) 2014 yarışmasında 22 katmandan oluşan GoogLeNet modeli %5,7 hata oranı ile birinci olmuştur. GoogLeNet, bellek ve güç kullanımını etkili bir biçimde kullanma konusunda da oldukça iyi bir başarıya sahiptir. Çok sayıda filtre ekleyerek katmanları yığnlamak, ek parametre hesaplamaları yapmayı, fazla bellek tüketimini ve ezberleme olasılığını artırmaktadır. GoogLeNet bu sorunların üstesinden gelebilmek için modülleri birbirine paralel bir biçimde bağlamaktadır (İnik ve Ülker, 2017). Aşağıdaki şekilde GoogleNet mimarisi gösterilmektedir.

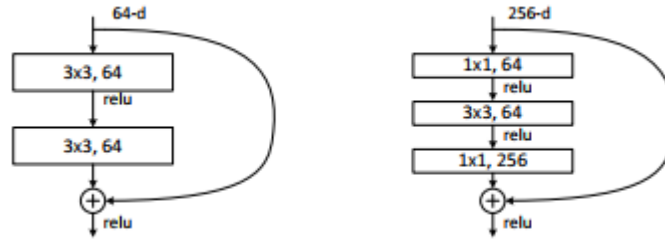


Şekil 3.16. GoogleNet mimarisi

Kaynak: (Öznel, 2018).

3.4.5 ResNet

2015 yılındaki ImageNet yarışmasının birincisi Zhang ve ark. (He vd., 2016) tarafından geliştirilen Microsoft ResNet mimarisi olmuştur. İnsanların nesne tanımda hata oranı %5-%10 olarak gösterilmekte iken bu mimari, %3,6 hata oranı ile nesne tanımda insan kabiliyetinin de üstüne çıkmaktadır (Öznel, 2019). ResNet, Şekil 3.17'deki gibi residual bloklardan oluşmaktadır. Bu bloklar kullanılarak bir katman tarafından üretilen çıktı aşağıda gösterildiği gibi daha sonraki katmanlara aktarılarak filtreleme işlemlerine devam etmektedir (Akın, 2019).



Şekil 3.17. Residual blok fonksiyonu

Kaynak: (Akın, 2019).

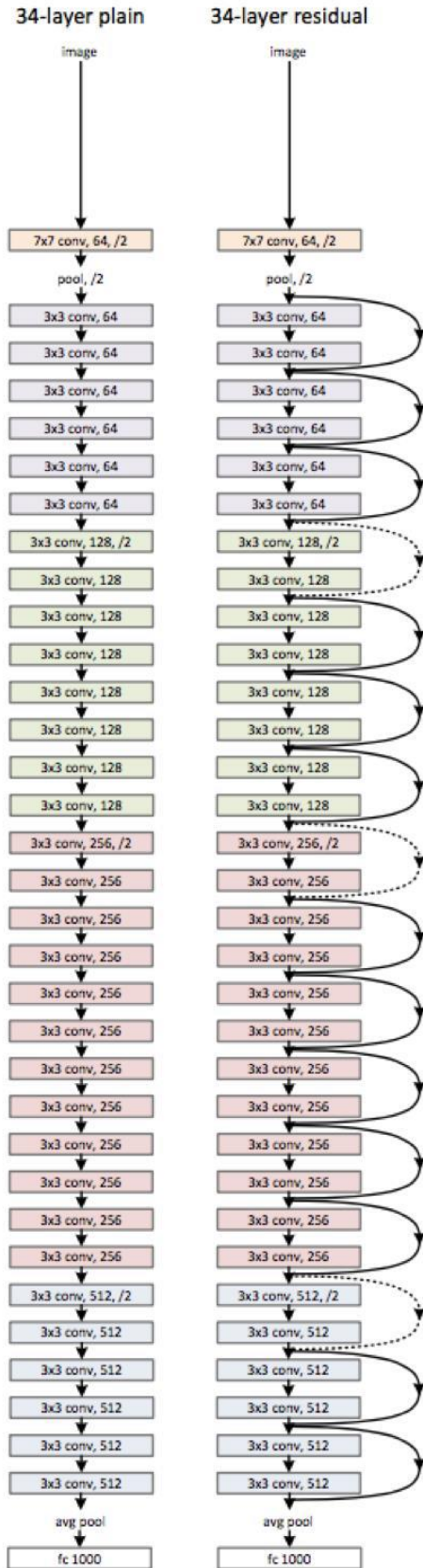
ResNet mimarisi, derin ağlardaki zorlaşan eğitim süreçleri için çözüm sağlamaya çalışmaktadır. Derin ağların, diğer yöntemlere göre daha başarılı sonuçlar verdiği belli olduktan sonra araştırmacılar bu modelleri daha verimli kullanmaya yönelmişlerdir. Ağın çok katmandan oluşması ve dolayısıyla çok daha karmaşık parametre hesabı yapılması bilinen bir gerçektir. Ancak, ilk başta bir ESA'yı daha derin tasarlanmanın sadece parametre sayısının artmasıyla oluşacak olan işlem yükünü olumsuz etkileyeceği düşünülürken ağın bir noktadan sonra artık öğrenemediği problemi tespit edilmiştir (Söylemez, 2021).

YSA'larda olduğu gibi ESA'lar da geri yayılım algoritmasıyla eğitilen modellerdir. Geri yayılım esnasında, hata fonksiyonunun çıktısı yardımıyla ağ üzerindeki ağırlıklar tekrar hesaplanmaktadır. Bu hesaplama esnasında zincir kuralından faydalanılır. Üst katmanlara doğru gidildikçe gradyan değerleri zincir kuralındaki çarpma teriminden dolayı azalmaktadır. Çok derin ağlarda bu değer sıfıra doğru yakınsamaktadır. Yani, sıfıra yakınsayan bu ağırlıklar öğrenmeye yeni bir katkı sunmamaktadırlar. Ortaya çıkan bu duruma kaybolan gradyan problemi

denilmektedir. ResNet, kaybolan gradyan sorununu çözmek için residual bağlantıları adı verilen bir yöntem önermektedir (He vd., 2016; Söylemez, 2021).

Geleneksel sinir ağlarında her bir katmanın çıktısı bir sonraki katmana girdi olarak verilmesine rağmen, residual bağlantılar ile bir katmanın çıktısı kendisinden sonraki katmana değil, daha sonraki katmanlara girdi olarak iletilebilmektedir. Bu metot ile iki tür yarar sağlanması hedeflenmektedir. Bunlardan birincisi, ağın ilk katmanlarından elde edilen özelliklerin çok fazla değişime uğramadan sonraki katmanlara iletilebilmesidir. İkincisi ise, geri yayılım esnasında bu bağlantılardan geçen gradyanların değişime uğramadan yukarıdaki katmanlara doğrudan iletilebilmesidir. Bu şekilde gradyanlar yukarıya doğru gittikçe büyüklükleri kontrol altında tutularak kaybolan gradyan sorununa çözüm sağlanmaktadır (He vd., 2016; Söylemez, 2021).

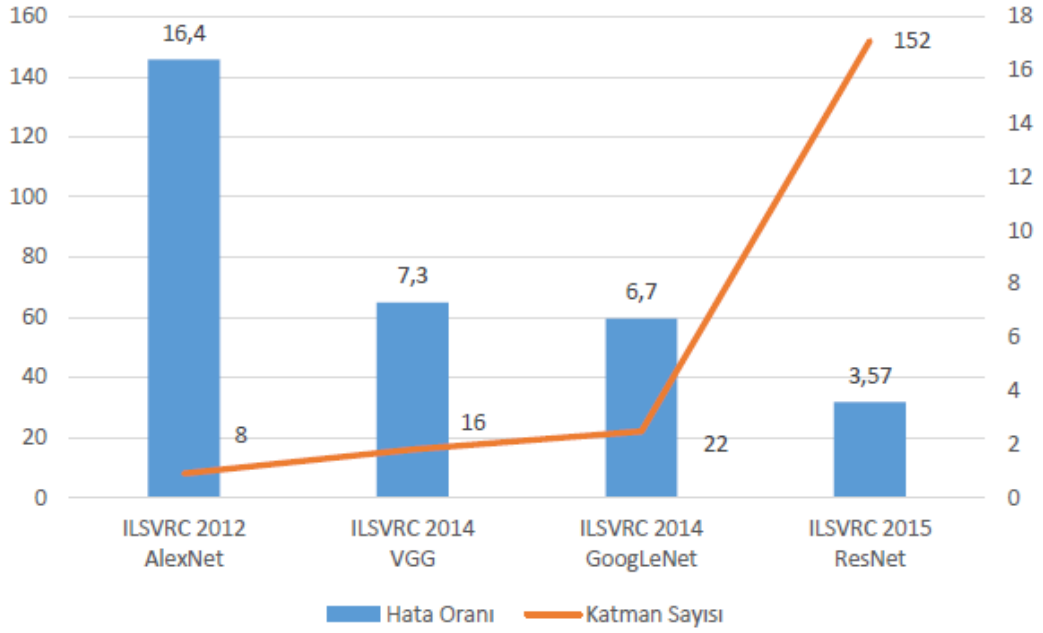
Şekil 3.18’de 152 katmandan oluşan Resnet mimarisi gösterilmiştir.



Şekil 3.18. ResNet mimarisi

Kaynak: (Özgel G. , 2019).

Buraya kadar bahsedilen ESA mimarilerinin katman sayıları ve ILSVRC yarışmasındaki ortaya koydukları başarı Şekil 3.19’da toplu olarak gösterilmektedir. Buna göre, ESA mimarileri sayesinde hata oranlarının giderek azaldığı söylenebilir. Dikkat edilmesi gereken bir diğer husus da grafikte görüldüğü üzere katman sayısı arttıkça sınıflandırma başarı oranının artmış olmasıdır (Akın, 2019).



Şekil 3.19. Bazı ESA mimarilerinin başarı karşılaştırması

Kaynak: (Akın, 2019).

3.5 Veri Artırma

Bir derin öğrenme modelinin daha iyi genelleştirme yapabilmesini sağlamanın en iyi yolu, daha fazla veri kullanarak modeli çok iyi bir biçimde eğitmektir. Ancak pratikte sahip olduğumuz veri miktarı sınırlıdır. Bu problemi aşmanın bir yolu, yapay veriler oluşturup bu verileri eğitim kümesine dâhil etmektir. Bu yaklaşım sınıflandırma için en kolay olan yöntemdir (Kutlugün vd., 2019). Bir sınıflandırıcının görevi, karmaşık, yüksek boyutlu bir x girdisini alıp bunu tek bir y kategorisi ile özetlemektir. Yani, bir sınıflandırıcının esas işi, çok çeşitli dönüşümlere karşı kararlı çalışmasıdır. Eğitim kümesindeki x girdileri dönüştürülerek kolayca yeni (x,y) ikilileri oluşturulabilmektedir (Goodfellow vd., 2016). Bu sayede veri çeşitliliği yapay yöntemlerle artırılmakta ve öğrenme başarımını büyük ölçüde etkilemektedir.

Veri artırma (data augmentation) yöntemleri ile görüntüler üzerinde bazı bozulmalar meydana getirilmektedir. Bu bozulmalar, açısız olarak görüntüyü döndürme, perspektifini değiştirme, belli bölgelere kaydırma, yakınlaştırma ve uzaklaştırma gibi farklı filtrelerin uygulanmasıyla görüntülerden çeşitli yapay görüntüler elde edecek şekilde gerçekleştirilmektedir (Şahin, 2017).

Şekil 3.20’de solda gözükten orijinal görüntü çeşitli filtreler kullanılarak sağdaki gibi yapay olarak artırılmıştır.



Şekil 3.20. Artırılmış görüntü örneği

Kaynak: (Gandhi, 2021).

Görüntüyü döndürmek veya ölçeklemek gibi birçok işlemin eğitim başarımını artırmada etkili olduğu gösterilmiştir. Ancak doğru sınıfı, yanlış olarak algılayabilecek dönüşümler yapmaktan kaçınmak gerekmektedir. Örneğin, bir optik karakter tanıma sisteminde “b” ile “d” harfleri veya “6” ile “9” rakamları arasında yanlış seçim yapmamak için yatay öteleme veya 180° açıyla döndürme gibi veri artırma yöntemleri uygun olmayabilir (Goodfellow vd., 2016; Kutlugün vd., 2019).

Kullanılan veri kümesi türü ve uygulama amacına bağlı olarak veriyi uygun olmayan biçimde artırmak, başarımlar açısından olumlu sonuçlar vermeyebilir. Veri kümesini artırmak için kullanılan filtreler ve bu filtrelerin uygun ölçülerde olması, daha ayırt edici özellikler elde etmek için oldukça önemlidir (Kutlugün vd., 2019).

Pek çok bilgisayarla görme görevi için veri boyutunu artırmak, eğitim veri kümesini genişleterek modelin farklı koşulları da öğrenmesine katkı sunmakta, ezberleme

durumu veya aşırı uydurma olarak tanımlanan overfitting'i önlemek için önemli bir yaklaşımdır (Zheng vd., 2019).

3.5.1 Veri Artırma Üzerine Yapılan Araştırmalar

Veri artırma yöntemleri birçok alanda ve veri kümesinde yaygın olarak kullanılmakta ve başarıya önemli ölçüde katkıda bulunmaktadır. Şahin Ö. yaptığı çalışmada veri büyütme tekniklerini kullanarak kullandığı görüntü sayısını 1540'tan 2640'a çıkarmıştır. 90 derecelik açısız döndürme, kaydırma ve yakınlaştırma filtrelerini farklı modellere uygulayarak, banknot veri setindeki farklı konumları elde etmek için ResNet50 modelinin eğitim verileri üzerindeki tutarlılığını %100 olarak hesaplamıştır (Şahin, 2017). Duman ve Akın çalışmalarında kullandıkları veri setini eğitime göre 1000, 3000 ve 20000 oranlarında arttırmışlar ve test başarısını %4'ten %57'ye çıkarmışlardır (Duman ve Akın, 2019). Krishevsky ve diğerleri tasarladıkları AlexNet modelinde aşırı uydurma problemini ortadan kaldırmak için veri artırma ve bırakma adı verilen iki teknik kullanmışlar ve sinir ağlarının derinliğinin ve eğitim aşamasında kullanılan veri miktarının sınıflandırma performansı için önemli olduğunu belirtmişlerdir. (Krizhevsky vd., 2012). Arpacı ve Varlı, U-Net modeli ile IOSTAR veri seti görüntülerinde mixup yönteminin retina damar bölütlemesine etkisini incelemişlerdir. Bu doğrultuda yatay aynalama, kırpma, ters çevirme gibi geleneksel veri artırma işlemlerine ek olarak mixup yönteminin uygulandığı farklı varyasyonları test etmişlerdir. Bu çalışmaya göre sadece Mixup'a uygulanan testlerde düşük performans elde edilirken, geleneksel ve mixup yöntemlerinin birlikte kullanıldığı test sonuçları daha yüksek başarı oranına sahip olmuştur (Arpacı ve Varlı, 2022).

Yüz tanıma çalışmalarında özellikle gerçek zamanlı sistemler için birçok veri artırma çalışması geliştirilmiştir. Genel olarak bu çalışmalar, poz açısı değişikliklerini tanıma ve yüz bölgesinde tanınmayı zorlaştıran etkenlere rağmen tanıma olarak sıralanabilir. Jiang-Jing Lv ve ark. çalışmalarında, CASIA-WebFace veri setini 25.580 denek ve 2.545.659 yüz görüntüsü içeren neredeyse iki kat daha büyük bir veri setine yükselterek, veri büyütmeyle dayalı büyütme ve manuel veri seti genişletme arasındaki performansı beş farklı yöntemde karşılaştırmışlardır. Son olarak, tüm bu modelleri birleştirerek %99,33 eğitim ve %94,08 test başarısı elde etmişlerdir (Lv vd., 2017).

Artan veri miktarıyla birlikte görüntü işleme alanında özellikle görüntü sınıflandırma, nesne tanıma gibi problemler öne çıkmaktadır. Artık büyük veri ile başa çıkabilecek bir yöntem olarak kendisini ispatlayan derin öğrenme metotları sıklıkla kullanılmakta ve başarılı sonuçlar vermektedir (Kutlugün vd., 2019).

Ancak, kısıtlı veri kümesine sahip uygulama alanları için verinin bütünlüğünü bozmadan veri miktarının artırılması önem arz etmektedir. Bu konuda Si Liu ve ark. (Liu vd., 2015), yapmış oldukları çalışmada kırılmış insan merkezli görüntü bölgelerindeki hassasiyet oranını artırmak ve aşırı uyumu azaltmak için görüntü boyutunu 1 ve 1,2 kat oranında büyütme, yatay yansıtma gibi filtreler ile her görüntünün 4 varyasyonu olacak şekilde artırmışlardır. Ancak bu çalışma sadece görüntü bölgesini içerecek şekilde yapılmış, yüz ayırma işlemleri gerçekleştirilmemiştir (Kutlugün vd., 2019).

Vittorio ve ark., çalışmalarında görüntü üzerindeki ayırt edici özellikleri zenginleştirmek için geniş bir artırılmış görüntü koleksiyonu üreterek önerdikleri metot ile aydınlatma, poz, yüz ifadesi ve düşük çözünürlük gibi kısıtları ortadan kaldırmaya çalışmışlardır (Cuculo vd., 2019). Buslaev ve ark., çalışmalarında görüntü artırma yöntemi ile çok sayıda görüntü dönüştürme işlemi yaparak hızlı ve etkili bir albüm oluşturma işlemi gerçekleştirmişlerdir (Buslaev vd., 2018).

Salman, çalışmasında sınıflandırma başarımını artırmak için mevcut eğitim veri kümesi üzerinden farklı seviyelerde gürültü eklemek suretiyle yapay veri kümesi üretimi gerçekleştirmiştir (Salman, 2018). Doğan ve Türkoğlu, derin öğrenme ile yaprak sınıflandırması için tasarladıkları modelde Alexnet, Vgg16, Vgg19, ResNet50, GoogleNet algoritmalarını kullanarak başarımlarını karşılaştırmışlardır. Buna göre, veri miktarının artması ile başarımın da artacağını ancak işlem süresinin de buna bağlı olarak uzayacağını belirtmişlerdir (Doğan ve Türkoğlu, 2018).

3.6 Boyut Azaltma

Boyut azaltma başarımı ve test sürelerini etkileyen bir adımdır. Boyut azaltma işlemleri ile kabul edilebilir işlem sürelerinde başarımı artıran sonuçlar elde etmek mümkündür (Kutlugün, 2017; Kutlugün ve Şirin, 2017). Bunun için istatistiksel analiz metotları, geniş ölçekli verilerin işlenmesinde yaygın kullanım alanına

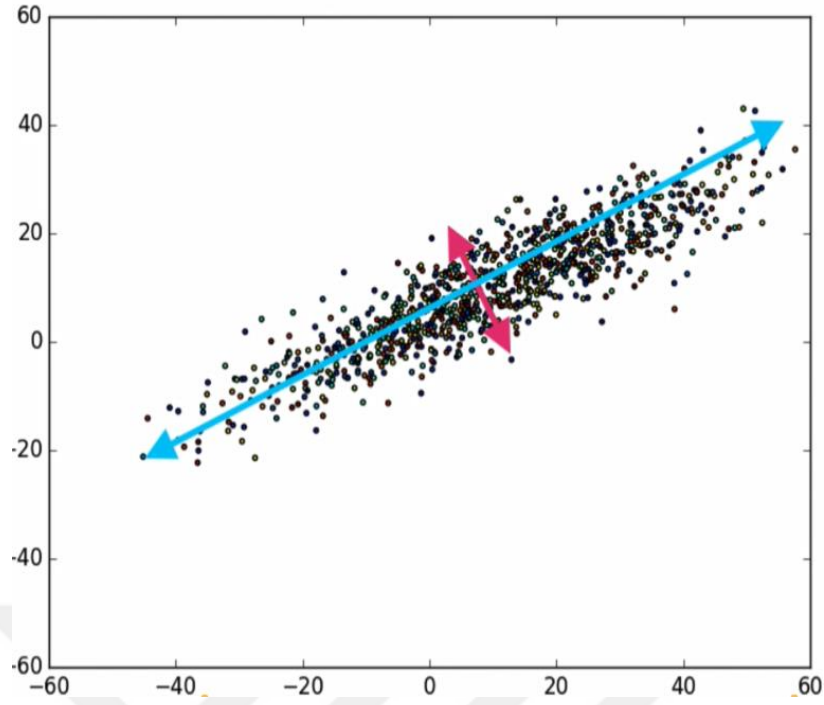
sahiptir. Öznitelik uzayı içinde örnekler arasındaki ilişkiyi en iyi şekilde ortaya çıkarabilme yeteneği bizlere etkili bir biçimde analiz yapma olanağı sağlamaktadır (Kutlugün ve Şirin, 2018b). İstatistiksel olarak ele alınan yöntemler sayısal çok güçlü bir alt yapıya sahip olduklarından çeşitli algoritmik araçlar ile verilerin daha etkili bir şekilde analiz edilebilmesine olanak sunmaktadırlar (Çalık, 2019).

Görüntü işlemede, boyut azaltmak için görüntü pikselleri yüksek boyutlu bir vektör uzayından düşük boyutlu bir öznitelik uzayına dönüştürülmeye çalışılmaktadır (Kutlugün, 2017; Kutlugün ve Şirin, 2018b). Kısaca, mevcut verilerin lineer birleşimleri ile yeni bir özellik uzayı elde edilmektedir. Bu amaçla, genellikle Temel Bileşen Analizi (Principal Component Analysis-PCA) (Moor, 1981) tercih edilmektedir. Ayrıca, Doğrusal Diskriminant Analizi (Linear Discriminant Analysis-LDA) (Balakrishnama ve Ganapathiraju, 1998), PCA'den daha güçlü bir özellik çıkarma ve tanıma tekniği olarak kullanılmaktadır. Son zamanlarda, t-Dağıtılmış Stokastik Komşu Yerleştirme (t-SNE), boyut azaltma için kullanılan daha belirgin bir algoritma haline gelmiştir (Van der Maaten ve Hinton, 2008).

3.6.1 PCA

Tanıma işleminde yapılacak en basit sınıflandırma yöntemi, test görüntüsünün eğitim kümesindeki en yakın komşusunun bulunması olarak söylenebilir. Bu sınıflandırma işlemi için veri kümesindeki görüntü boyutu kadar işlem yapılmalı ve görüntüler arasındaki ilişkiye dikkat edilmelidir. Tüm bu işlemler hesaplama maliyetini artırmaya sebep olmaktadır. Temel Bileşen Analizi (TBA) ile tanıma görüntüsünün boyutunun azaltılması ve yeni bir formda ifade edilmesi gerçekleştirilmektedir (Moor, 1981; Yazıcı, 2008).

Temel Bileşen Analizi (PCA), sınıf bilgisini dikkate almayan denetimsiz (unsupervised) bir yöntemdir. Büyük boyuttaki verileri çok fazla kayba uğratmadan daha düşük boyuttaki veriler ile temsil edebilmeyi amaçlamaktadır. PCA, uygulama kolaylığı sağlaması sebebiyle yüz tanıma uygulamalarında sıkça kullanılmaktadır (Şahin, 2020). TBA vektörlerinin oluşumunu gösteren temsili bir gösterim Şekil 3.21'de sunulmuştur.

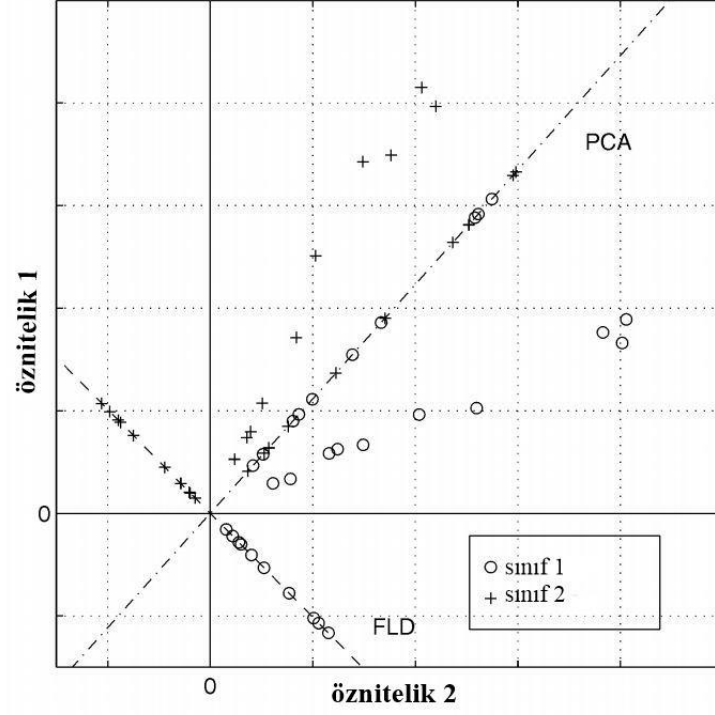


Şekil 3.21. PCA vektörleri

3.6.2 LDA

LDA, sınıf odaklı doğrusal yaklaşım sunarak PCA ve diğer yöntemlere göre daha fazla avantaj sağlamaktadır. LDA yöntemi, bir yüz tanıma probleminde doğrusal karar sınırlarından yararlanarak farklı aydınlatma koşulunda özellik vektörü elde etmeye çalışır (Balakrishnama ve Ganapathiraju, 1998; Karaman, 2019).

PCA, verilerin sınıflarını göz ardı ederek boyut indirgeme işlemini gerçekleştirmektedir. Bundan dolayı sınıf örneklerinin iç içe girebildiği örnek durumlarda sınıflandırma başarımı düşeceği için PCA tercih edilmemektedir. Bunun üstesinden gelebilmek için LDA yöntemi önerilmiştir. LDA, varyans değerlerinin yanında sınıf bilgisini de dikkate almaktadır. Kısaca LDA, denetimli (supervised) boyut indirgeme işlemi yapmaktadır. Şekil 3.22’de LDA yöntemi ile öznelilikler arasında sınıfsal bağlantı kurarak indirgeme yapan bir örnek gösterim sunulmuştur.



Şekil 3.22. LDA vektör ve sınıfları

Kaynak: (Karaman, 2019).

3.6.3 T-SNE

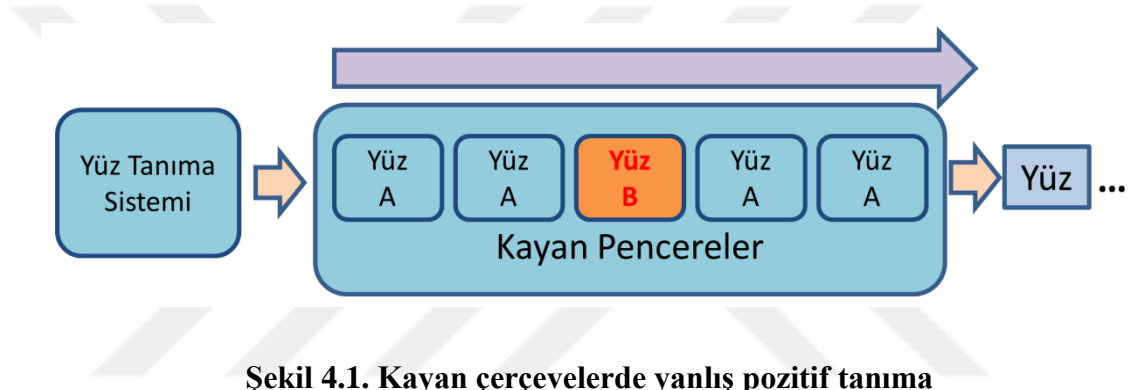
T-SNE yönteminin amacı, sınıflar arasındaki uzaklıkları mümkün olduğu kadar koruyarak daha düşük boyutlu temsiller elde edebilmektir. Öncelikle, her bir veri noktasının rastgele elde edilmiş düşük boyutlu temsilleri ile işleme başlanır. Daha sonra örneklem uzayındaki yakın olan noktaları daha yakın, uzak olanları ise daha uzak tutacak şekilde işlemlere devam edilir. Uzak olan noktaların uzaklığını korumak yerine yakın noktaları daha çok önemser (Van der Maaten ve Hinton, 2008; Koptur, 2017). Son zamanlarda, t-SNE, yüksek boyutlu verilerin daha düşük boyutlarda uygun temsilleri nedeniyle en popüler boyut indirgeme algoritmalarından biri haline gelmiştir.

DÖRDÜNCÜ BÖLÜM

DENEYSEL ÇALIŞMA ve UYGULAMA GERÇEKLEME

4.1 Deneysel Çalışmanın Amacı

Gerçek zamanlı yüz tanıma sistemlerinde anlık görüntü akışına en uygun modelin yüksek başarımla çalışması beklenmektedir. Bu anlık değişen karelerde zaman zaman Şekil 4.1’de gösterildiği gibi bazı hatalı tanımlama veya kayıplar meydana gelebilmektedir.



Şekil 4.1. Kayan çerçevelerde yanlış pozitif tanıma

Kaynak: (Sütçüler, 2006).

Meydana gelen bu hatalı tanımlamaların çok çeşitli sebepleri olabilir. Bu sebeplerden bazıları görüntülerde poz açısı veya ışık şiddeti gibi unsurların tanımayı zorlaştıracak veya hatalı tanıma yol açacak kadar değişikliğe uğramasıdır. Özellikle, eğitim veri kümesinde bulunan az sayıda örnek içeren görüntüler bulunması, bu hatalı tanımlamalara sebep olabilmektedir. Bu çalışmanın amacı, yüz tanıma süreci ve gerçek zamanlı yüz tanıma sistemlerinde karşılaşılan hatalı tanımlamaların olası sebeplerini açıklamak ve bu durumların azaltılması için çözüm sağlayabilecek modeller önermektir. Bunun için üç farklı uygulama gerçekleştirilmiştir. Bu uygulamalara ait süreçler kendi alanları içerisinde aşağıda detaylı bir biçimde açıklanmaktadır.

4.2 Uygulama-1: Veri Artırma Yöntemleri ile Aşırı Uyum Probleminin Ele Alınması

Yüz tanıma sistemlerinde ışık yönü, yansıma, yüz ifadesindeki duygusal ve fiziksel değişiklikler tanımayı zorlaştıran temel etkenlerdir. Sistemin küçük veri kümelerinde az sayıda numuneden oluşan verilerle eğitilmesi, performansı olumsuz etkileyen önemli bir faktördür. Oysaki Evrimsel Sinir Ağı (ESA) temelli modeller, büyük miktarda eğitim verisi için saklı özellikler ile en iyi özellik haritasını çıkarmada kullanılan derin öğrenme mimarileridir. Bu uygulamanın amacı, kaynak veri kümesindeki az sayıda örnek içeren yüz görüntülerinden, orijinal görüntünün bütünlüğünü bozmadan, çok sayıda örnek ile tasarlanan ESA modelinde kullanılmak üzere eğitim verileri elde etmektir. Bu sayede gerçek zamanlı olmayan yüz tanıma sisteminin sınıflandırma başarımını nasıl etkilediği gözlemlenecektir.

4.2.1 Veri Kümesi (Dataset)

Veri kümesi için, her kişiye ait sadece bir adet vesikalık fotoğraf görüntülerinden oluşan 150 adet küçük bir veri kümesi kullanılmıştır. Verinin gizliliği ilkesi gereği bu görüntüler paylaşılmamış, sonuçlar bir tabloda sunulmuştur. Ayrıca, Literatürde Labeled Faces in the Wild (LFW) (Huang, 2008) olarak bilinen veri kümesinden ilave edilen rasgele seçilmiş 150 kişiye ait görüntüler ile test yapılmıştır.

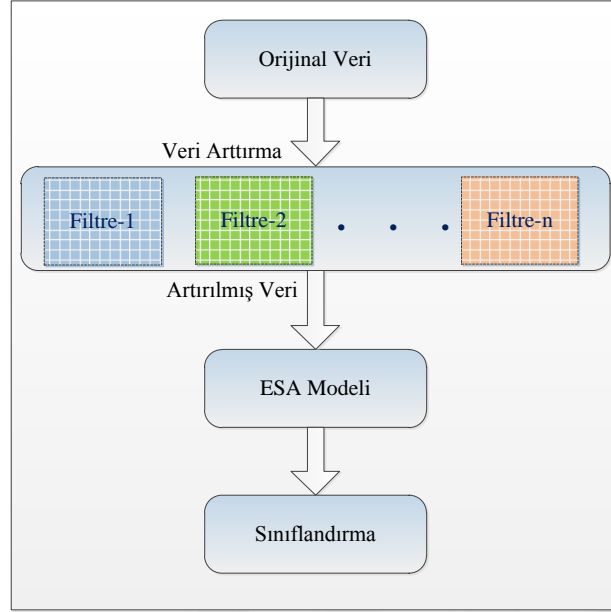
4.2.2 Veri Önleme

Kullanılan farklı boyutlardaki tüm görüntüler 256x256 boyutuna getirilmiş, giriş katmanında yer alacak olan görüntü boyutu 256x256x3 olarak ele alınmıştır.

4.2.3 Uygulama Gerçekleme

Uygulama, Python platformunda Keras Kütüphanesi kullanılarak gerçekleştirilmiştir. Veri artırmak için Keras'ın ImageDataGenerator parametreleri kullanılmıştır. Uygulamaya ait süreç diyagramı Şekil 4.2'de gösterilmiştir.

Buna göre veri artırma tekniklerinden faydalanılarak her yüz görüntüsü için 22, 42 ve 63 kat filtreleme ile görüntü sayısı kademeli olarak 150'den sırasıyla 3300, 6300 ve en son 9450'ye çıkarılmıştır. Bu veri kümelerine; açısal döndürme, kaydırma, kırpma, yakınlaştırma, çevirme ve parlaklık fitreleri uygulanmış, ortaya çıkan boşlukları doldurmak için k-en yakın komşu algoritması tercih edilmiştir.



Şekil 4.2. Süreç Blok Diyagramı

Uygulanan filtreler ile elde edilen artırılmış görüntüler Şekil 4.3'te gösterilmektedir. Şekilde sol üst köşede gösterilen görüntü orijinal görüntü, diğerleri ise artırılmış görüntülerdir.



Şekil 4.3. Veri artırma ile elde edilen görüntüler.

LFW veri setinden rastgele seçilen 150 kişiye ait yüz görüntüsünü karşılaştırmak için herhangi bir veri artırma işlemi uygulanmamıştır. Çünkü LFW veri seti içerisinde zaten bir kişiye ait birden fazla yüz görüntüsü örnekleri bulunabilmektedir.

Uygulamanın başarımı, eğitim veri kümesinin doğruluk yüzdesine göre ölçülmüştür. Sadece bir adet yüz görüntüsünden oluşan orijinal veri kümesinin veri artırma

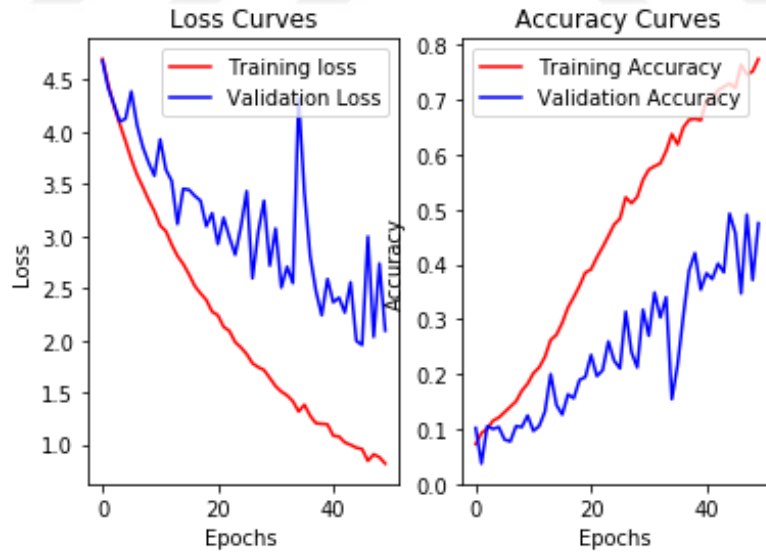
öncesinde ve sonrasında elde edilen toplam veri miktarı ile ortalama başarımların yüzdeleri aşağıdaki tabloda sunulmuştur.

Tablo 4.1. Veri artırma başarımların değerlendirme tablosu

Olay	Veri Sayısı	Eğitim Süresi	Başarımların (~%)
1	150	58 sn.	50
2	3300	20 dk.	62
3	6300	44 dk.	66
4	9450	1s. 4 sn.	76

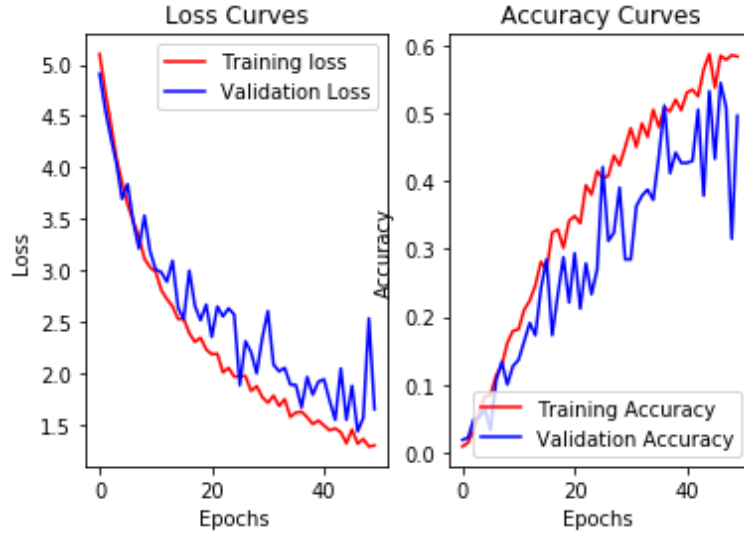
4.2.4 Uygulama Değerlendirme

Şekil 4.4'te görülen grafik, veri artırma işlemi uygulanmadan önce, yani sistemin ilk haliyle test edilmesi sonucu oluşmuştur. Şekilden de görülebildiği gibi oldukça verimsiz bir sonuç alındığı net olarak söylenebilir. Bu haliyle sistemi daha fazla eğitmek ve böylece başarımların yüzdesini artırmaya çalışmak oldukça uzun zaman alacaktır. Öncelikli olarak verilerin çeşitliliğini sağlamaya çalışmak daha uygun bir çözüm sunabilir.

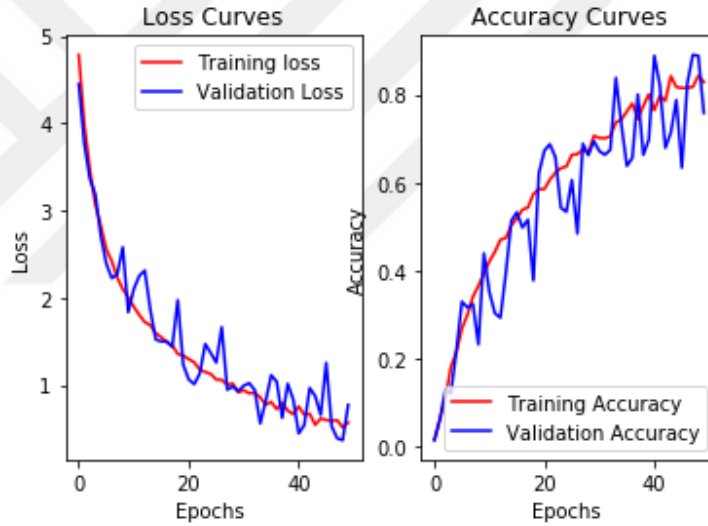


Şekil 4.4. Veri artırma uygulanmadan önce elde edilen eğitim grafiği.

İkinci olarak, 6300 ve 9450 adete kadar artırılmış veri kümelerinin eğitim doğruluk-kayıp grafikleri Şekil 4.5 ve Şekil 4.6'da gösterilmektedir. Bu grafiklerden de görüleceği üzere veri artırma işleminin eğitim başarımlarına doğrudan katkı sağladığı ve makul eğitim sürelerinde bu oranlara ulaşılabilirdiği söylenebilir.



Şekil 4.5. 6300 adet veri artırma uygulanarak elde edilen eğitim grafiği.



Şekil 4.6. 9450 adet veri artırma uygulanarak elde edilen eğitim grafiği.

Bu çalışmada kullanılan filtrelerden özellikle açısız döndürme ve parlaklık filtrelerinin başarıma daha fazla olumlu etkisi olduğu görülmüştür. Özetle, deneysel sonuçlar, veri miktarına bağlı olarak eğitim başarımının gözle görülür seviyede arttığını göstermektedir. Eğitim başarımını daha fazla artırmak için uygun filtreler makul parametre değerleri ile yeniden tasarlanabilir veya eğitim işlemi daha uzun sürecek şekilde gerçekleştirilebilir. Ancak, çok yüksek veri miktarı ile ortaya çıkabilecek ezberleme veya aşırı uyum problemi ile veri miktarına bağlı artan eğitim süresi belli bir seviyeden sonra başarıma daha fazla olumlu katkı sunmayabilir.

4.3 Uygulama-2: Yüz Tanıma Sistemlerinde Kademeli Benzerlik Oranları Yöntemi ile Eğitim Veri tabanının Artırılması

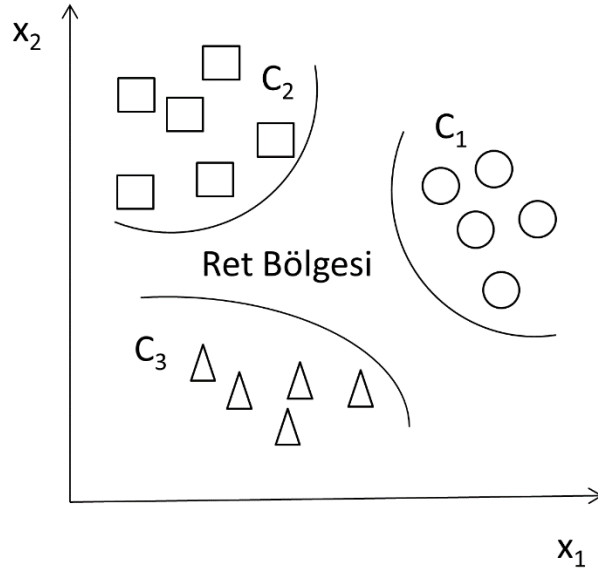
4.3.1 Amaç

Temsili öğrenme adı verilen derin metrik öğrenme algoritmaları yüz tanıma sistemlerinde sıklıkla tercih edilmektedir. Ancak modelin öznelik çıkarımındaki başarısının yanı sıra, bu veri tabanındaki örneklerin dağılımı ve uygun sınıflandırıcı tercihleri gibi faktörler de yüz tanıma sisteminin genel performansını etkilemektedir.

Bu çalışma, önceden eğitilmiş bir derin sinir ağı modelinden oluşturulacak veri tabanında, sınırlı sayıda numune içeren sınıflar için nitelikli artırılmış veriler elde ederek bütünlük oluşturmayı amaçlamaktadır. Kademeli Benzerlik Oranları (Gradual Similarity Ratios, GSR) (Kutlugün ve Şirin, 2023) adı verilen bu yöntem sayesinde sınıf bütünlüğünü bozabilecek artırılmış numuneler veri tabanından kaldırılmıştır. Bu sayede sınırlı örneğe sahip sınıflar bir bütünlük içinde tutulmakta ve sınıflandırıcı davranışı daha etkin kullanılmaktadır. Özetle, çalışmanın katkıları aşağıdaki gibi maddeler halinde sunulabilir:

- sınırlı verilerle sınıfları en iyi şekilde artırma
- aşırı veri artırmayı filtreleme
- sınıflandırıcıyı daha verimli kullanma
- küçük boyutlu vektörlerde bile yüksek performans sağlama

K-nn sınıflandırıcılar başta olmak üzere derin metrik öğrenme tabanlı yüz tanıma sistemlerinden elde edilen 128 boyutlu vektör özneliklerinin sınıf bütünlüğünün korunması esastır. Çünkü iç içe girmiş şekilde dağılan veriler yanlış pozitif tanımaya sebep olabilecektir. Bu nedenle, verileri artırırken hedeflenen sınıftan uzaklaşma eğiliminde olabilecek aşırı artırılmış numuneler filtrelenmeli veya hariç tutulmalıdır. GSR, bu mesafeyi belirlemek için yeni bir yöntem sunarak sınıflandırıcı eşik değerinden yararlanılabileceğini önermektedir. Bu şekilde, sınıflara ait yalnızca gerçekten ayırt edici özellikler içeren numuneler saklanacak ve veri tabanında bütünlük oluşacaktır. Hedeflenen dağılım, Şekil 4.7'de temsili olarak sunulmuştur. GSR ile filtrelemenin detayları Bölüm 4.3.3'te detaylı olarak anlatılmıştır.

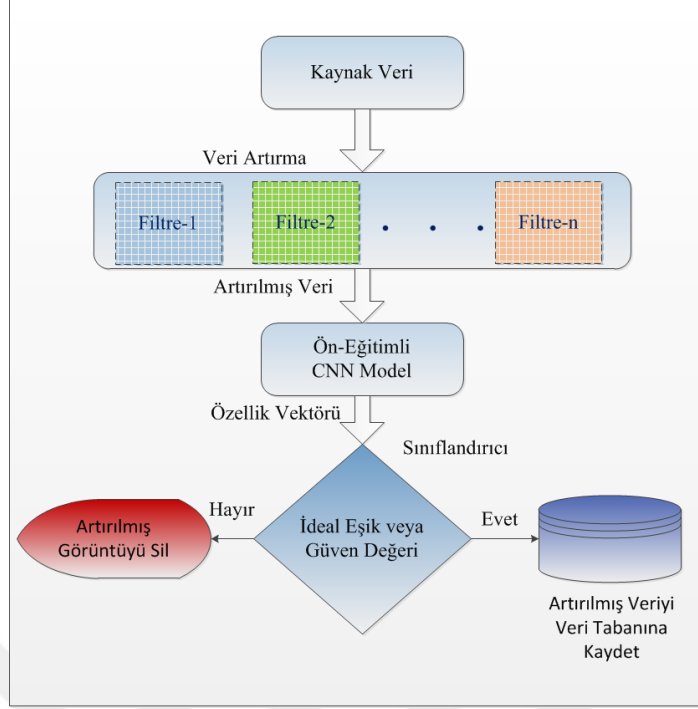


Şekil 4.7. İdeal sınıf ayrımı ile veri dağılımı.

4.3.2 Model ve Mimari

Çalışmada model olarak Dlib yüz tanıma kütüphanesi kullanılmıştır. Dlib kitaplığına göre varsayılan sınıflandırıcı eşik değeri 0,6 olarak belirlenmiştir ve daha düşük oranlar, yüz karşılaştırmalarını daha katı veya hassas hale getirmektedir (King, 2017). Bu oran ile LFW veri kümesinde en yüksek başarı oranına ulaşıldığı için uygulamada bu değer varsayılan temel değer olarak kabul edilmiştir. Farklı hassasiyet istenen durumlar için bu oran isteğe bağlı olarak değişiklik gösterebilir. Bu durum önerilen GSR yönteminin ana işlevini etkilemeyecektir.

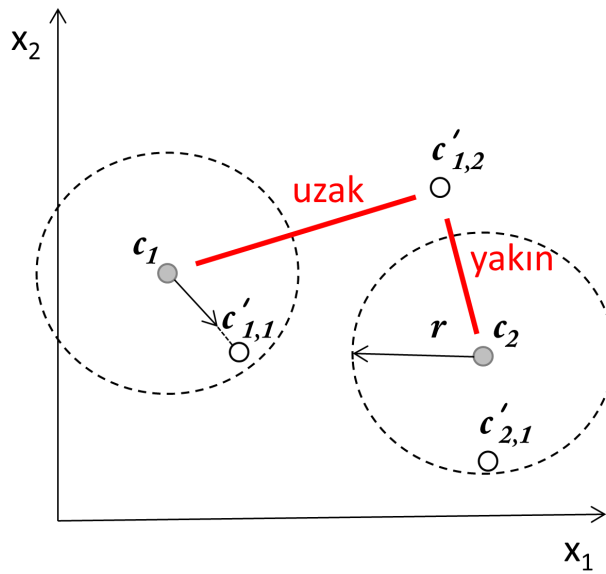
Modele göre Şekil 4.8'deki akış şemasında gösterildiği gibi, bazı filtreler uygulanarak orijinal görüntü sayısı artırılmaktadır. Arttırılmış her bir görüntü, Konvolüsyonel Sinir Ağları (CNN) modeli ile 128 boyutlu vektörlere dönüştürülür. GSR bu aşamadan sonra devreye girerek sınıflandırıcının eşik değerini göz önünde tutarak filtreleme işlemi yapmaktadır. Sınıflandırıcı aşamasında, yalnızca kabul edilen eşik değeri sınırına uygun olan artırılmış numuneler veri tabanında tutularak diğer artırılmış yapay numuneler veri kümesine dâhil edilmeyip atılacaktır.



Şekil 4.8. Mimari.

4.3.3 Kademeli Benzerlik Oranları (Gradual Similarity Ratios, GSR)

c_1 ve c_2 , sınırlı sayıda veriye sahip iki farklı sınıfın elemanları olsun. Veri büyütme yöntemleriyle c_1 sınıfından türetilen ilk numuneye $c'_{1,1}$, ikincisine $c'_{1,2}$ ve c_2 sınıfından türetilen ilk örneğe $c'_{2,1}$ diyelim. Rasgele veri artırma işlemleri sonucunda Şekil 4.9'daki gibi veri dağılımı oluştuğunu varsayalım.



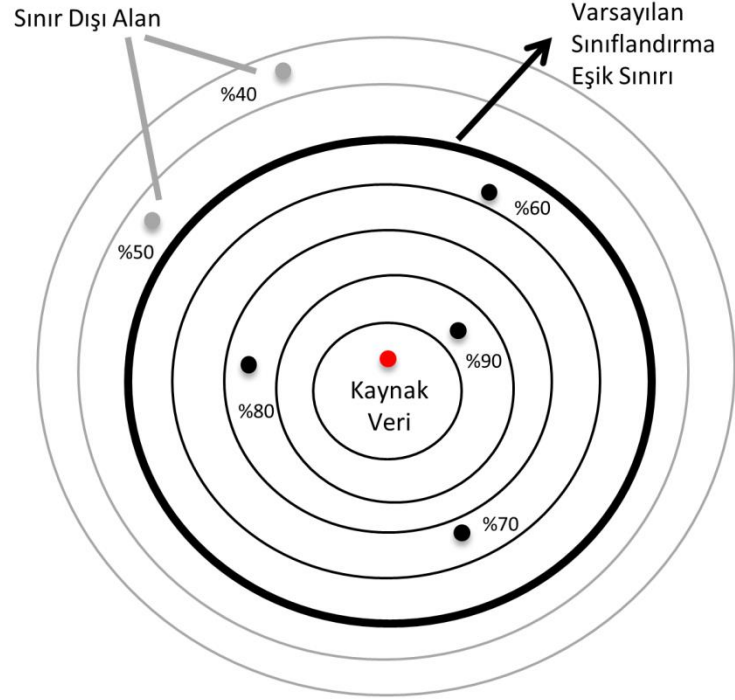
Şekil 4.9. GSR eylemi ile eşik ilişkisi.

Bu durumda, elde edilen her bir yeni numune veri tabanında depolanırsa, c_1 'den türetilen ikinci numune, c_2 sınıfına daha yakın olduğu için yanlış bir şekilde yanlış pozitif olarak sınıflandırılmış olacaktır. Burada $c'_{1,2}$ numunesi için aşırı veri büyütme ile aşırı bozulmaya maruz kaldığı söylenebilir. Böylece kendi sınıfından uzaklaşp başka bir sınıfa yaklaşarak dengesiz bir dağılıma neden olmuştur. Bu tür problemlerin önüne geçebilmek için önerilen GSR yöntemi ile sınıflandırıcı eşik değeri dikkate alınarak veri miktarının yapay olarak artırılması kontrol altında tutulmaktadır. Buna göre denklem (1)'deki Öklid uzaklık formülü (Danielsson, 1980) ile iki vektör karşılaştırıldığında, sınıflandırıcı eşik değerini aşan numuneler denklem (2) kuralına göre filtrelenerek elenecektir. Bunun yerine sınıflandırıcı eşik değerini aşmayan $c'_{1,1}$ ve $c'_{2,1}$ numuneleri sınıfın uygun alanlarında saklanacaktır.

$$d(x, x') = \sqrt{\sum_{i=1}^N (x_i - x'_i)^2} \quad (1)$$

$$GSR \text{ Action} = \begin{cases} 0, \text{discard} & r < d(x, x') \\ 1, \text{accept} & r \geq d(x, x') \end{cases} \quad (2)$$

GSR'nin ikinci bir işlevi ise, eşik sınırı içindeki veriler için her farklı yüzdelik alana karşılık gelen bölümleri ayrı ayrı filtrelemektir. Veri büyütme filtresiyle orijinal bir görüntüden altı adet sentetik veri elde etmek istediğimizi varsayalım. Bu durumda Şekil 4.10'daki gibi GSR kullanarak veri artırma yapılabilir. Ancak %50 ve %40'lık dilimlere karşılık gelen kukla verilerin sınıflandırıcı eşiklerinin dışında kalmasından dolayı kullanılamayacağına dikkat edilmelidir. Diğer yüzdelik dilimlerde, kalan dört yapay veri yüzdelik dilim başına bir veriyle sınırlandırılmıştır. Bu sayede daha homojen bir veri dağılımı sağlanmaktadır. Aksi halde dört yapay verinin orijinal veriye en yakın konuma konumlandırılması sayı olarak uygun olmakla birlikte nitelik olarak uygun olmayacaktır. Çünkü veri çoğaltmanın temel amacı, farklı durumları öğrenmek ve veri çeşitliliğini sağlayarak bunları doğru bir şekilde sınıflandırmaktır. Kısaca, her yüzdelik dilimden eşit sayıda yapay numune oluşturmak veri bütünlüğü sağlayacaktır. Ancak burada farklı filtreler uygulayarak aynı yüzdelik dilimlere birden çok sayıda veri daha da çeşitlendirilebilir. Örneğin, kırpma filtresi ile dört farklı yüzdelik dilime karşılık gelen yapay numune oluşturduktan sonra, aydınlatma filtresi ile de yine dört farklı yüzdelik dilime yeni yapay veriler ilave edilebilir.



Şekil 4.10. Kademeli benzerlik oranlarına göre dağıtılan veriler.

4.3.4 Uygulama

Özellikle sınırlı sayıda örnek içeren farklı veri kümeleri ile deneysel çalışmalar yapmak için birçok farklı veri artırma yöntemleri uygulanmıştır. Bu uygulamalar sonucunda modelden elde edilen benzerlik oranlarına göre GSR eylemi incelenerek önerilen yöntemin etkisi ölçülmüştür.

Yüz tanıma çalışmalarında sıklıkla kullanılan LFW veri kümesi 13000'den fazla yüz görüntüsü içermektedir. Bunlardan 1680 kişinin iki veya daha fazla yüz görüntüsü mevcuttur. Uygulamada bunların dışında kalan ve sadece bir örneği bulunan 4069 kişinin yüz görüntüleri kullanılmaktadır. Orijinal verilere 150 kişilik gruplar halinde rasgele seçilen sınırlı alt örneklem kümeleri ile veri artırma filtreleri uygulanmıştır.

İlk aşamada, veri artırma teknikleri kullanılarak her bir yüz görüntüsü için açısız döndürme, kaydırma, kırpma, yakınlaştırma ve aydınlatma filtreleri gibi geleneksel yöntemler uygulanmıştır. Böylece Şekil 4.11'de en soldaki orijinal görüntü artırılarak yapay iki numune görüntü elde edilmiş ve GSR yöntemi ile filtreleme uygulayarak uygun olan bir numuneyi veri tabanına kabul etmiş, diğerini eşik değerinin dışına çıkarttığı için ret etmiştir.

Kaynak Görüntü	Artırılmış Görüntüler	
		
GSR Eylemi:	Benzerlik Oranı: 0,76 Kabul	Benzerlik Oranı: 0,51 Ret




Şekil 4.11. Geleneksel yöntemler uygulanarak artırılmış görüntüler.

İkinci aşamada, geleneksel veri artırma yöntemlerinin yanı sıra, farklı sınıflar arasındaki görüntülere son teknoloji MixUp (Zhang vd., 2017) yöntemi (alfa:0,2 ile) uygulanmıştır. MixUp işlemi sonucu oluşturulan yeni görüntüler, GSR eylemi ile birlikte Şekil 4.12’de sunulmuştur.

a) Kaynak Yüz	b) Farklı Yüz	c) MixUp ile elde edilen Yüz	
			Benzerlik Oranı: 0,53 GSR Eylemi: Ret

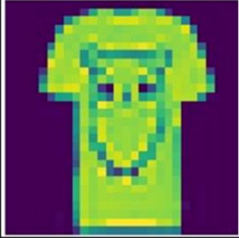
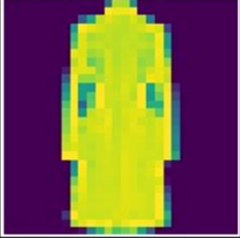
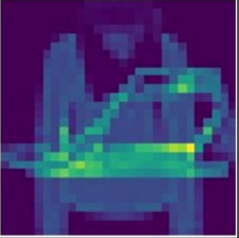
Şekil 4.12. Farklı sınıflar arasında uygulanan MixUp görüntüleri.

Üçüncü aşamada, aynı sınıfa ait orijinal görüntüye farklı bir veri artırma işlemi uygulanmış ve ardından MixUp yöntemiyle (alfa:0,2 değerinde) yeni görüntüler elde edilmiştir. Bu işlemlerden elde edilen yeni numuneler Şekil 4.13’te GSR eylemi ile birlikte gösterilmiştir.

a) Kaynak Yüz	b) Yaşlandırma Filtresi ile Yüz	c) MixUp ile Yüz Görüntüsü	
			Benzerlik Oranı: 0,99 GSR Eylemi: Kabul




Şekil 4.13. Aynı sınıfa uygulanan MixUp görüntüleri.

Dördüncü aşamada, MixUp yöntemi yüz veri tabanları dışında başka bir veri kümesi için uygulanmıştır. Bu uygulama sonucunda FashionMNIST veri kümesinden (Xiao, 2017) elde edilen görüntüler GSR eylemi ile Şekil 4.14'teki gibi gösterilmiştir.

MixUp Görüntüler		
		
Benzerlik Oranı: 0,96 Kabul	Benzerlik Oranı: 0,82 Kabul	Benzerlik Oranı: 0,52 Ret

Şekil 4.14. FashionMNIST veri kümesi için MixUp örneği.

Son aşamada, CutMix (Sangdoon vd., 2019) yöntemi iki farklı sınıfa uygulanmıştır. CutMix işlemi ile oluşturulan yeni görüntü ve GSR eylemi Şekil 4.15'teki gibi gözükmemektedir.

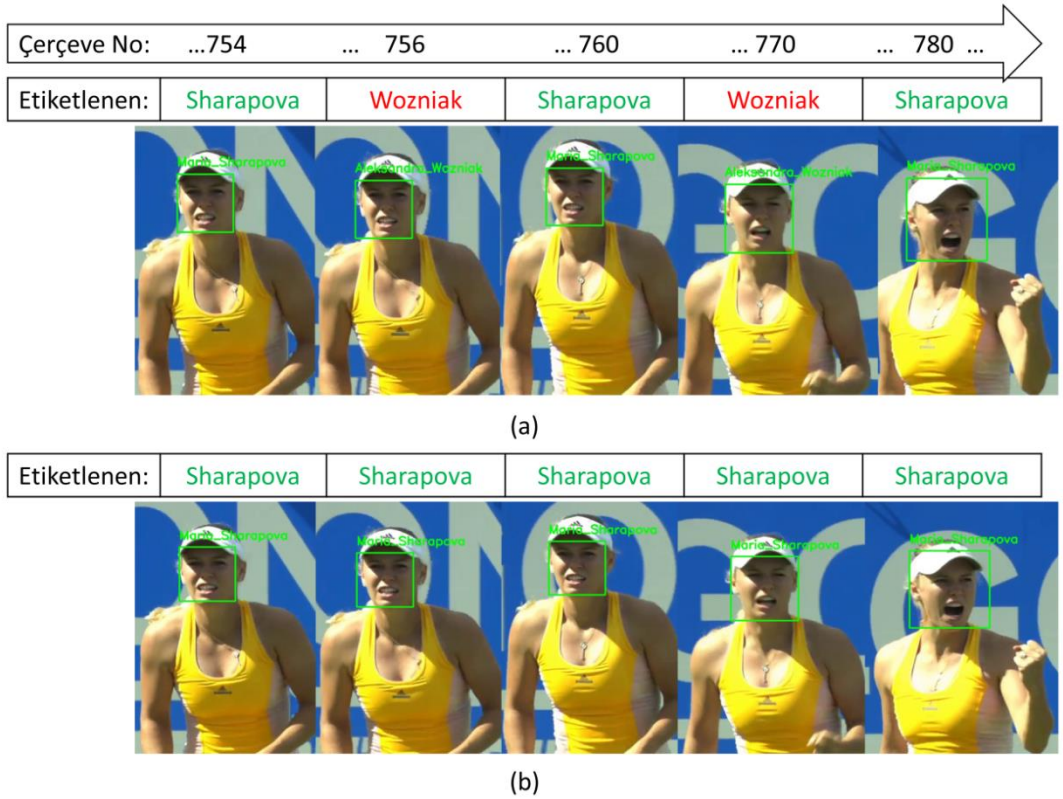
a) Kaynak Yüz	b) Farklı Yüz	c) CutMix ile elde edilen Yüz	
			Benzerlik Oranı: 0,51 GSR Eylemi: Ret

Şekil 4.15. Farklı sınıflar arasında uygulanan CutMix yöntemi ile elde edilen görüntüler.

Ayrıca GSR'nin gerçek zamanlı yüz tanıma sistemi üzerindeki etkisini ölçmek için Tablo 4.2'de detaylandırılan test videosu ile karşılaştırma yapılmıştır. İlk olarak, mevcut veri kümesine Aleksandra Wozniak ve Maria Sharapova'nın görüntüleri eklenerek rastgele veri artırımı yapılmıştır. Elde edilen 516 adet yapay görüntü önce Şekil 4.16(a)'daki gibi herhangi bir filtreleme yapılmadan (GSR uygulanmadan) karşılaştırılmış, daha sonra GSR yöntemi ile 25 adet numune içerecek şekilde süzülüş ve Şekil 4.16(b)'deki sonuçlar elde edilmiştir.

Tablo 4.2. Test videosunun özellikleri

Video Etiketi	“Wozniak ve Sharapova”
Süre (s)	324
Çerçeve Sayısı	8103
FPS	25
Boyut	854x480
Kategori	Spor-Tenis
Video Özellikleri	Çeyrek finalde farklı oyuncuların hareketli kamerayla çekilmiş özet maç görüntülerini içerir. Youtube-8M dataset orijinal başlık: “2015 Aegon International Eastbourne Quarterfinal WTA Highlights”



Şekil 4.16. GSR ile veri artırma test videosunda yüz tanıma performansının karşılaştırılması.

4.3.5 Uygulama Sonuçları ve Değerlendirme

Bu değerlendirme sürecinde denklem (3)'teki Kesinlik (Olson ve Delen, 2008) (P) metriği kullanılmıştır.

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (3)$$

Şekil 4.16(a)'da rastgele veri artırma yöntemlerinden elde edilen yapay veriler, sınıflar arasındaki mesafenin oldukça yakın olduğunu göstermektedir. Bu nedenle, iki sınıfın elemanları karıştırılabilir ve farklı yüz pozlama açılarında kayıplar meydana gelebilir. Bu test videosunda geleneksel yöntemle yaklaşık %80 doğruluk değeri elde edilmiştir.

Şekil 4.16(b)'de yapay veriler GSR ile filtrelendiği için daha kararlı sonuçlar elde edilebilmektedir. Bu test videosunda GSR yöntemi ile yaklaşık %88 doğruluğa ulaşılmıştır.

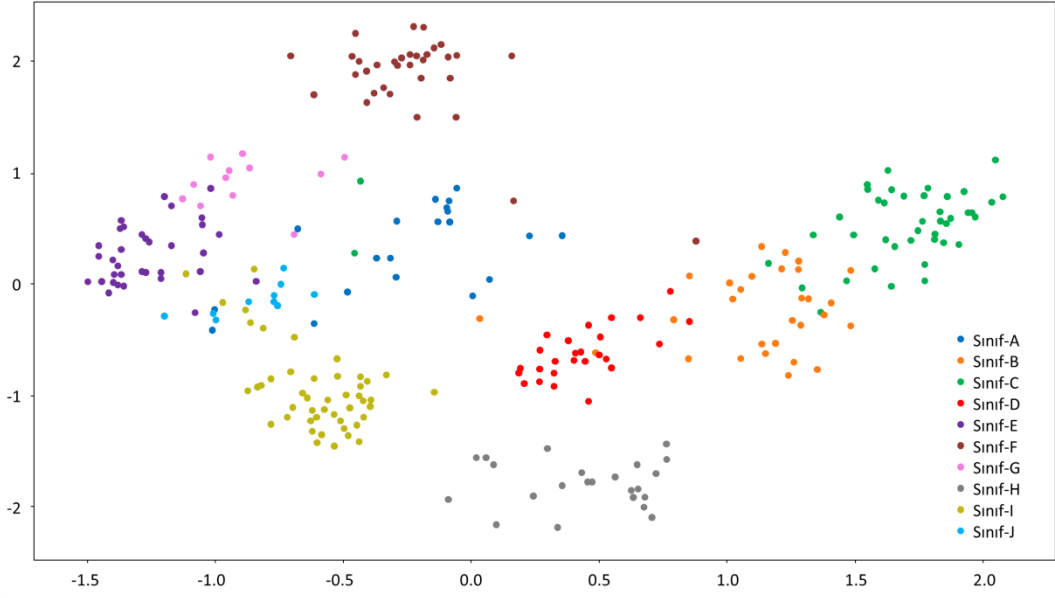
Ayrıca, farklı yöntem ve seviyelerde gerçekleştirilen uygulamalardan elde edilen sonuçlar Tablo 4.3'te sunulmuştur.

Tablo 4.3. Farklı veri artırma yöntemlerinin 3 seviyeli performans tablosu

Yöntem	Veri Miktarı		
	Seviye - 1	Seviye - 2	Seviye - 3
GSR içermeyen Geleneksel	3300	6300	9450
Başarım Oranı (%)	62.38%	66.56%	76.84%
Geleneksel + GSR	2400	5250	6150
Başarım Oranı (%)	93.74%	96.36%	98.82%
Geleneksel + Sınıf bazlı MixUP + GSR	2250	4650	5700
Başarım Oranı (%)	94.9%	97.86%	99.38%

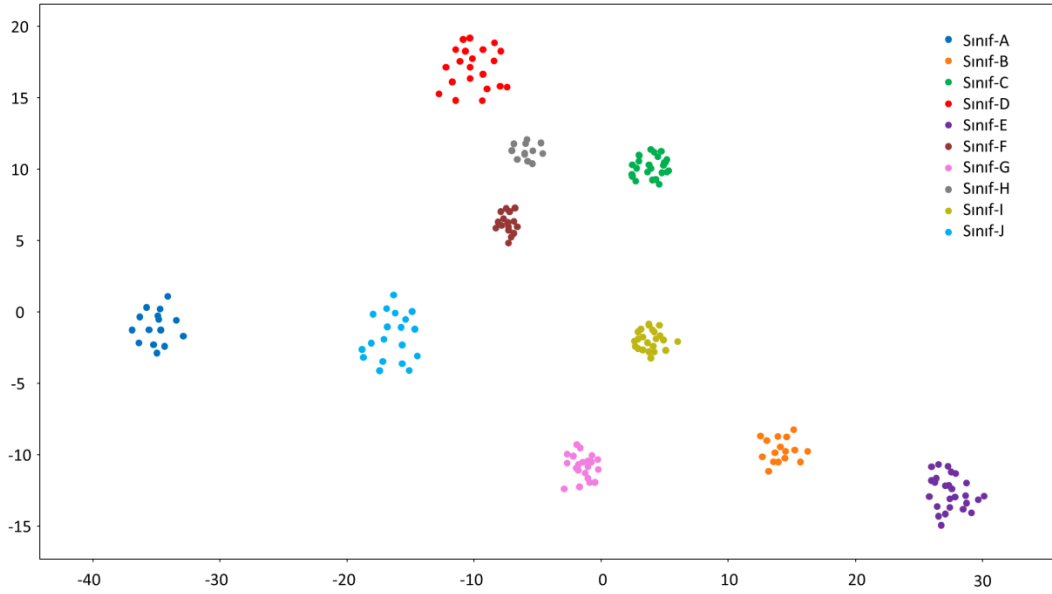
Kaynak: (Kutlugün, M.A. and Şirin, Y., 2023).

Geleneksel veri artırma yöntemleri ile rastgele aşırı oranlarda artırılarak oluşturulan veri tabanının dağılımı Şekil 4.17'deki gibi oluşabilir. Bu yöntem ile şekildeki verilerin yaklaşık %77 doğrulukla dağıldığı söylenebilir.



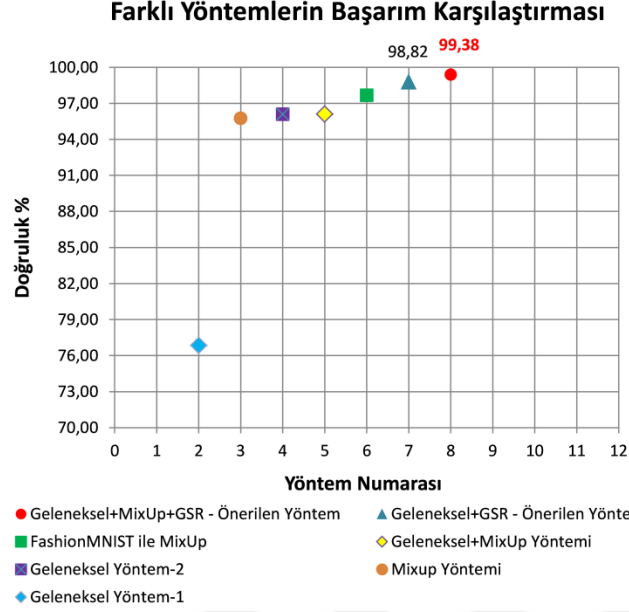
Şekil 4.17. Rastgele artırılmış eğitim veri kümesi dağılımı.

GSR yöntemi ile eşik değeri dışındaki verilerin atılması sonucunda elde edilen veri kümesi dağılımı temsili olarak Şekil 4.18'deki gibi oluşabilecektir. Buna göre verilerin yaklaşık %99 doğrulukla dağıldığı görülmektedir.



Şekil 4.18. GSR uygulanan veri kümesi dağılımı.

Yöntem-1 (Kutlugün vd., 2019) ve Yöntem-2 (Arpacı ve Varlı, 2022) gibi farklı uygulama ve farklı veri kümeleri ile yapılan karşılaştırmalardan elde edilen performans dağılım grafiği Şekil 4.19'daki gibidir.



Şekil 4.19. Farklı yöntemlerin başarımlarının karşılaştırılması.

Ayrıca, 128 boyutlu varsayılan vektör büyüklüklerinin t-SNE (Van der Maaten ve Hinton, 2008) yöntemi ile farklı boyutlara indirgenmesi sonucu oluşan başarımların oranları Tablo 4.4'te sunulmuştur. Bu durum, daha düşük vektör boyutlarında bile kabul edilebilir düzeyde başarımlarına sahip olduğunu göstermektedir.

Tablo 4.4. Boyut küçültme etkileri tablosu

Vektör Boyutu	t-SNE (%)
64	98.75
32	98.52
16	98.29
8	98.13
4	97.91
2	97.78

Sonuç olarak, bu çalışmada önerilen yöntem ile veri tabanının sistematik bir düzende şekillendirilmesi sağlanmıştır. Bu kademelendirme, veri kümesinde çok sayıda düzensiz veriye sahip olmak yerine, uygun ölçütlere sahip etkin veri örnekleri içermesini sağlayacaktır. Veri kümesinde kabul edilebilir sınırlar içinde her seviyeden farklı numuneler içermesi sistemin çok daha tutarlı davranmasını sağlayabilecektir. Bu sayede sınırlı örneğe sahip sınıflar bir bütünlük içinde tutulmakta ve sınıflandırıcı davranışı daha etkin kullanılabilecektir.

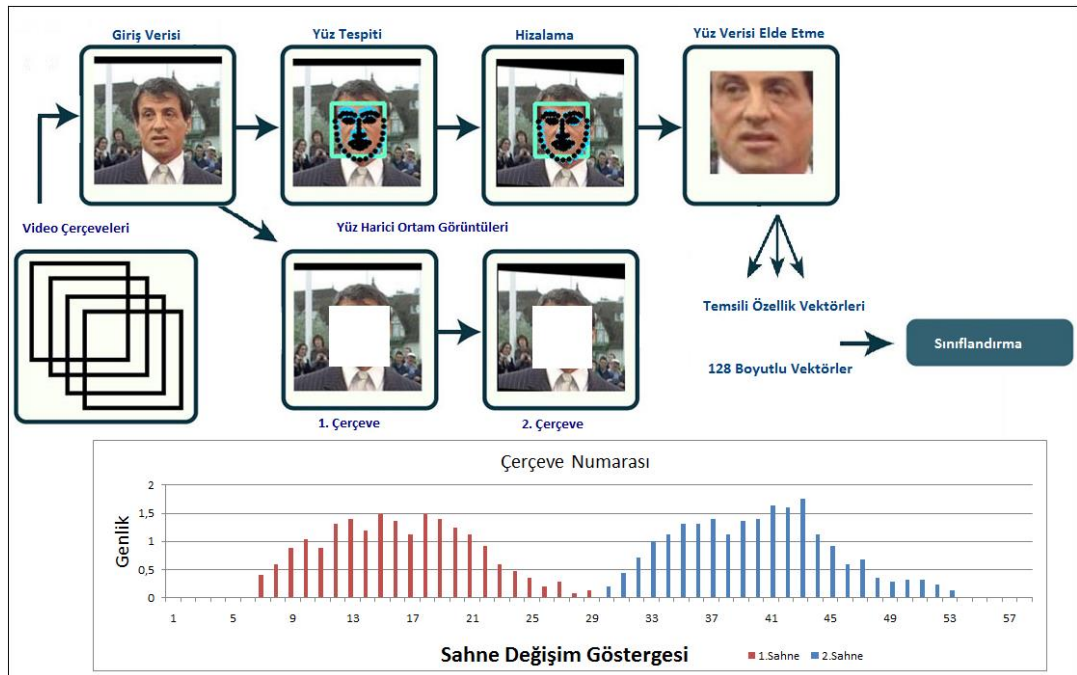
Deneysel alıřmada nerilen model, geleneksel veri artırma modellerine gre %99,38 doęruluk deęerlerine ulařmıřtır. Deneysel sonular, veri tabanının daha kk vektr boyutlarında bile kabul edilebilir bir bařarı dzeyine sahip olduęunu ve daha dzenli bir yapıda olabileceęini gstermektedir.

Ancak, eřik oranının uygun deęerlerde belirlenmedięi durumlarda yz tanıma uygulaması sonucunda “Bilinmeyen kiři” etiketi ile karřılařılması olasılıęı ortaya ıkabilecektir. Bu durum, tanınması istenen hedef grntnn kendi sınıf sınırlarının dıřına ıkararak ret blgesine denk dřebileceęi anlamına gelmektedir.

Veri kmesinin daęılımına ek olarak, zellikle k-nn sınıflandırıcıları kullanan modellerde, veri artırma kapasitesi (veri tabanında bulunacak her sınıftaki numune sayısı) de nemli bir etkendir.

4.4 Uygulama-3: Arka Plan deęişimlerinin Gerçek Zamanlı Görüntüler Üzerinde Bir Parametre Olarak Kullanılması

Gerçek zamanlı yüz tanıma sistemlerinde genel olarak, bir video görüntüsündeki belli bir süre boyunca akan veya kayan pencereler olarak adlandırılan çerçeveler ele alınarak işleme başlanmaktadır. Sahne üzerinde bulunan kişilere ait yüz tespiti işlemi sürecin başlangıcındaki en önemli aşamadır. Tespit edilen yüz görüntüleri bir model ile daha önceden eğitilerek veri tabanında saklanan özellikler olarak bazı sayısal değerlere dönüştürülmektedir. Elde edilen bu değerler, veri kümesinde kimliklendirilerek saklanmakta ve belli ölçütlerle karşılaştırılarak sınıflanmakta veya tanıma görevini yerine getirmektedir. Bu aşamaya kadar kayan pencereler üzerindeki sadece yüz bölgeleri ile işlem yapılmakta, yüz bölgeleri dışında kalan bölgeler herhangi bir deęer ifade etmemektedir. Önerilen bu çalışma ile birbirini takip eden her sahne görüntüsü içindeki ortam veya sahne deęişiklikleri de dikkate alınmakta ve gerçek zamanlı yüz tanıma sistemine katkı sağlanması amaçlanmaktadır. Böylece sahne ortamında belirgin düzeyde deęişmeler olmadıkça, yüz tanıma sistemi tarafından yanlış pozitif olarak nitelendirilebilecek hataların azaltılmasına katkı sunulması ve daha tutarlı bir yüz tanıma sistemi oluşturulması amaçlanmaktadır. Önerilen çalışmaya ait süreçler temsili olarak Şekil 4.20'deki gibi gösterilmiştir.

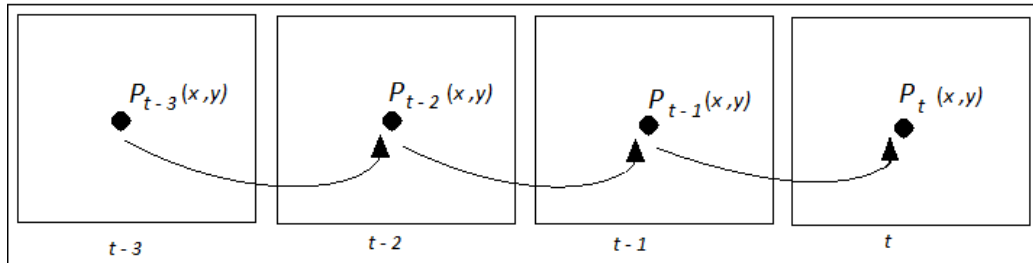


Şekil 4.20. Sahne deęişim göstergesi ile gerçek zamanlı yüz tanıma süreci.

4.4.1 Sahne Değişikliklerini Algılama Yöntemleri

Ani olarak meydana gelen sahne değişikliklerinde önceki sahnenin bitiminden hemen sonra yeni bir sahne başlayarak geçiş anındaki bazı çerçeveler arasındaki benzerlikler kayba uğramaktadır. Dereceli olarak değişen sahnelerde ise, ani geçişlerin aksine bir sahneden diğerine geçişler kademeli olarak gerçekleşmektedir. Önceki sahneye ait görüntü belli bir çerçeve süresince etkisini yavaşça azaltarak yeni sahneye ait görüntü yavaşça daha belirgin hale gelmektedir. Bundan dolayı kademeli sahne geçişlerindeki bir birini takip eden çerçeveler arasında yüksek benzerlikler korunmaktadır. Kademeli sahne değişiklikleri günümüzde uygulamada en çok kullanılan zincirleme sahne geçiştir (Güllü vd., 2004).

Video sahnelerindeki art arda geçiş yapan sahnelerin tespit edilebilmesi için piksel temelli, model temelli, hareket temelli ve histogram temelli bazı yöntemler mevcuttur. Bu ardışıl çerçevelerdeki piksel yoğunluk değerlerinin mutlak değişimi belli bir eşik değeri ile izlenerek çok fazla değişim özelliği gösteren, yani zincirleme sahne geçiş özelliği taşımayan pikseller ayırt edilebilmektedir. Zincirleme sahne geçişlerinin tespiti için belirlenen sayıda çerçeve izlenerek aşağıdaki şekildeki gibi piksel değerlerinin değişim bilgileri elde edilebilmektedir (Güllü vd., 2004).

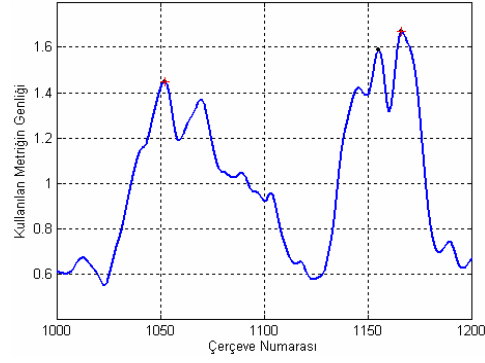


Şekil 4.21. Arka arkaya üç kez aynı değişim özelliğini gösteren piksellerin tespit edilmesi

Kaynak: (Güllü vd., 2004).

Belirlenen bir eşik değerinin üzerinde tespit edilen bölgelerde aşağıdaki şekilden de görülebileceği üzere birden fazla yerel tepe noktaları oluşabilir. Film içerisinde, otuz çerçevelik dilimlerdeki piksellerin artması, azalması ve değişmemesi gibi özellikler izlenerek bu özelliklerin oranı ile elde edilen metriğin yerel tepe noktalarına ulaşılabilir (Güllü vd., 2004).

Bu yöntem ile bir sahne değişim göstergesi oluşmaktadır.

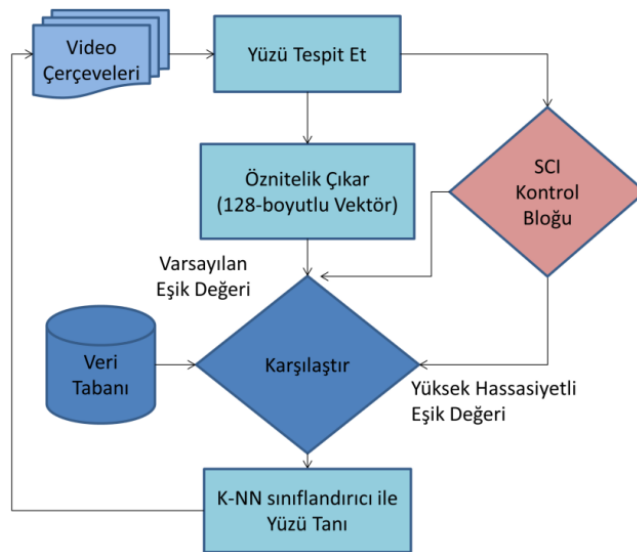


Şekil 4.22. Zincirleme sahne geçiş grafiği örneği

Kaynak: (Güllü vd., 2004).

4.4.2 Model

Yapılan uygulamada, yüz tespiti aşamasında Dalal ve Triggs (Dalal ve Triggs, 2005) tarafından 2005 yılında geliştirilen Histogram of Oriented Gradient (HOG) algoritması tercih edilmiştir. Bu algoritma, Viola-Jones (Viola, 2004) algoritmasına kıyasla aydınlatma, ölçekleme ve gölgeleme gibi pek çok farklı koşul altında daha hızlı ve doğru sonuçlar vermektedir. Yüz tespiti için, Davis King (King, 2017) tarafından oluşturulan Dlib modeli esas alınmıştır. Bu model, her bir yüz görüntüsü için 128 boyutlu vektör uzayı ile temsil edilebilen gömme (embedding) kullanmaktadır. Daha sonra bu vektörler k-en yakın komşu algoritmasına göre sınıflandırılmakta ve yüz tespiti gerçekleştirilmektedir. Önerilen modelin mimarisi aşağıdaki gibidir.



Şekil 4.23. Önerilen modelin mimarisi.

4.4.3 Sahne Değişim Göstergesi Bloğu (Scene Change Indicator, SCI)

Bu çalışmada SCI ile video görüntüler içindeki sahne geçişlerini bulmak için piksel yoğunluğu izlenerek tespit yapılmaya çalışılmıştır. Eşikleme ile belirli bir değere sahip piksel değerlerinin mutlak değişimi ve çok fazla değişiklik gösteren ve dolayısıyla sahne geçiş özelliği olmayan pikseller aşağıdaki gibi ayırt edilmektedir (Güllü vd., 2004).

$$D_{i,j}^p = |I_t(x, y) - I_{t-1}(x, y)| \quad (1)$$

$$M_t(x, y) = \begin{cases} 1, & D_{i,j}^p < th_m \\ 0, & other \end{cases} \quad (2)$$

$D_{i,j}^p$, (i,j)nci önceki ve mevcut bloklar arasındaki piksel düzeyi farkıdır. I_t ve I_{t-1} , video dizisinde sırasıyla t ve $(t-1)$ anlarında ardışık görüntü çerçevelerini temsil eder. Oluşturulan M_t maskesinde sahne geçişlerini algılamakta kullanılan piksellere ait bilgiler tutulmaktadır (Güllü vd., 2004). Burada kullanılan th_m eşik değerini temsil eder. Bu modül için piksel tabanlı sahne değişikliği tespiti yerine histogram tabanlı (Lee ve Cho, 2022) veya farklı bir detektör de kullanılabilir.

Kullanılan yöntemin algoritması Algoritma 1'de sunulmuştur.

Algoritmada, mevcut çerçeve ile önceki çerçeve arasındaki farklı blokların sayısını belirlemek için $D^{c,r}$ kullanılır. Tf_{imm} , bir sahne değişikliği olup olmadığını tespit etmek amacıyla karşılaştırmada kullanılır. Td_{block} , geçerli çerçeve ile referans alınan çerçeve arasındaki farklı blokların sayısını belirlemek için kullanılır. $D^{c,r}$, önceden tanımlanmış Tf_{imm} 'den büyükse, sahnenin durumu değişmiştir. Aksi takdirde, sahnenin değişmediği anlamına gelir.

Algoritma 1: SCI algoritması

```
Giriş: GirişVideoÇerçeve(V)
Çıkış: YüzTanıma
yüz_sayısı: 0

Prosedür SCI(V)
  Giriş Özellikleri(V)
  Başlangıç eşik:  $Td_{\text{blok}}$ ,  $Tf_{\text{imm}}$ , varsayılan_eşik, karşılaştırmaya_yeni_eşik
   $D^{c,r} = \text{SCD}(\text{mevcut\_çerçeve}, \text{önceki\_çerçeve}, Td_{\text{blok}})$ 

  if( $D^{c,r} < Tf_{\text{imm}}$ ) //sahne değişmedi
    YüzTanıma(mevcut_çerçeve, karşılaştırmaya_yeni_eşik)
  else if( $D^{c,r} > Tf_{\text{imm}}$ ) // sahne değişti
    YüzTanıma(mevcut_çerçeve, varsayılan_eşik)
  end if
Prosedür sonu

fonksiyon SCD(mevcut, referans,  $Td_{\text{blok}}$ )
  {  $B_{0,0}, B_{0,1}, B_{0,2}, \dots, B_{N-1}, B_{M-1}$  } = referans NxM blok
  for n=0 to N-1
    for m=0 to M-1
       $\Delta_{n,m}^{c,r} = \text{abs}(B_{n,m}^c - B_{n,m}^r)$ 
      Artır  $D^{c,r}$ , if  $\Delta_{n,m}^{c,r} > Td_{\text{blok}}$ 
    end for
  end for
  Çıktı  $D^{c,r}$ 
fonksiyon sonu

fonksiyon YüzTanıma(mevcut_çerçeve)
  [yüz_sayısı, sınırlayıcı_kutu] = YüzTespit(mevcut_çerçeve)
  yüz_kodlama = ÖzellikÇıkarma(sınırlayıcı_kutu)
  YüzKarşılaştırma(yüz_kodlama)
  sıfırla yüz_sayısı
fonksiyon sonu
```

4.4.4 Veri Kümesi (Dataset)

Modelin uygulama aşamasında Adile Naşit, Ata Demirer, Mehmet Öz, Mesut Özil ve Naim Süleymanoğlu ile Youtube-8M (Abu-El-Haija vd., 2016) veri kümesindeki Aleksandra Wozniak, Maria Sharapova, Martin Palermo, Matt Smith ve Morgan Freeman'dan oluşan ünlüler veri kümesi kullanılmıştır. Daha sonra bu kümeye LFW (Huang vd., 2008) veri setinden ve önceki uygulamada kullanılan yakın çevremizden

topladığımız bazı görüntüler de ilave edilerek yaklaşık 419 kişilik bir veri kümesi oluşturulmuştur. Bu 419 sınıfın her biri on adet farklı görüntü içermektedir. Veri kümesinin bir örneği Şekil 4.24'te gösterilmektedir.

Bu eğitilmiş veri kümesinden 128 boyutlu yüz vektörlerini (gömmeler) içeren veri tabanı oluşturularak, gerçek zamanlı test için yukarıda bahsedilen ünlülerin video görüntülerini içeren, sistemin daha önce görmediği 12 adet video seti ile test edilmiştir. Bu video kliplerin özellikleri Tablo 4.5'te gösterilmiştir. Bu test videolarında ön işleme yapılmıştır. İlk olarak, sahne geçiş yerlerinin ve sahnenin çok az değiştiği yerlerin manuel olarak belirlenmesi gerçekleştirilmiş, ikinci olarak özellikle sahnenin değişmediği kareler için önerilen model ile karşılaştırmalar yapılmıştır.



Şekil 4.24. Veri Kümesi örneği.

Tablo 4.5. Veri kümesinin özellikleri

Video No	Video klip adı	Süre (s)	Çerçeve sayısı	FPS	Boyut	Kategori	Video klbin özellikleri
1	"Adile Naşit"	58	892	25	854x468	Sinema / Tiyatro	Sanatçının başrolünü üstlendiği "Dream Family" filminden bir bölüm içermektedir. Bu bölümde farklı sahnelerde farklı kişiler gösteriliyor.
2	"Aleksandra Wozniak"	134	4028	30	1280x720	Spor - Tenis	Oyuncu, bir röportajda sabit kameraya açıklamalarda bulunuyor. Youtube-8M veri kümesindeki orijinal başlık: "WTA Tennis Star and USANA Brand Partner Aleksandra Wozniak talks about her life."
3	"Ata Demirel"	113	2843	25	640x360	Sinema/ Komedî	Sanatçının "One Man Giant Staff" isimli stand-up gösterisinde farklı kamera açılarıyla çekilmiş görüntüleri yer alıyor.
4	"Maria Sharapova"	246	6172	25	854x480	Spor - Tenis	Oyuncunun rakibiyle oynadığı maçın bir bölümünü içerir. Youtube-8M veri kümesindeki orijinal başlık: "2015 Internazionali BNL d'Italia Final WTA Highlights Maria Sharapova vs Carla Suarez Navarro"
5	"Martin Palermo"	127	3175	25	640x480	Spor - Futbol	Maçtan sonra stadyumda hareketli kamera ile kaydedilen oyuncunun veda konuşmasını içerir. Youtube-8M veri kümesindeki orijinal ad: "Palermo en la Bombonera, la despedida"
6	"Matt Smith"	140	3504	25	1280x720	Sinema / Tiyatro	Aktör Matt Smith'in BBC'nin Doctor Who dizisindeki sahneleri içerir. Youtube-8M veri kümesindeki orijinal adı: "Matt Smith Tribute Doctor Who HD."
7	"Mehmet Öz"-1	297	7444	25	640x360	Sağlık	Dr. Öz, farklı sahnelerle sabit bir kameraya Korona virüsün korunma planı hakkında bilgi veriyor.
8	"Mehmet Öz"-2	182	4555	25	640x360	Sağlık	Dr. Öz, dışarıda hareketli bir kamera ile röportaj yapıyor.
9	"Mesut Özil"	94	2352	25	640x360	Spor - Futbol	Oyuncu, maç sonrası bir reklam standının önünde canlı bir kamerayla röportaj yapıyor.
10	"Morgan Freeman"	161	3876	24	1280x720	Sinema	Aktörün oynadığı farklı sahneleri içeren bir hareketli kamerayla çekilmiş bir film klipinden görüntüler içerir. Youtube-8M veri kümesindeki orijinal başlık: "Gone with the Wind (6/10) Movie CLIP - To Lose a Child (2007) HD."
11	"Naim Süleymanoğlu"	65	1955	30	640x360	Spor - Halter	Oyuncunun Avrupa Şampiyonası için Cardiffteki halter müsabakasından sahneleri içerir.
12	"Wozniak ve Sharapova"	324	8103	25	854x480	Spor - Tenis	Çeyrek finalde farklı oyuncuların hareketli kamerayla çekilmiş özet maç görüntülerini içerir. Youtube-8M veri kümesindeki orijinal başlık: "2015 Aegon International Eastbourne Quarterfinal WTA Highlights"

4.4.5 Yöntem

Önerilen yöntem, sahne değişim hareketlerinin tanıma hassasiyeti üzerinde etkili olabileceği düşüncesi temel alınarak ortaya atılmıştır. Yeni bir parametre ile gerçek zamanlı video görüntülerinde sınıflandırıcı aşamasında sınıfın belirlenmesi için kullanılan benzerlik eşik değerinin dinamik olarak değiştirilmesi esasına göre çalışmaktadır. Bahsedilen parametrenin elde edilmesi için sahne değişim göstergesinden yararlanılmaktadır. Sahne değişim göstergesinden elde edilen bilgilere göre, değişikliğe uğramayan veya çok az değişime uğrayan sahne çerçeveleri belirlenmektedir. Bu çerçeveler arasında yüz görüntüsünün sınıfının atanması için kullanılmakta olan sabit karşılaştırma eşik değeri daha fazla hassasiyete sahip bir değere dönüştürülerek yüz tanıma işlemi gerçekleştirilmektedir. Muhtemel sahne değişikliği durumlarında Dlib modelinde en uygun değer olarak önerilen sabit değere (0,6) dönülerek en yüksek doğruluk oranına ulaşılması hedeflenmektedir.

4.4.6 Uygulamanın Gerçekleştirimi

Veri setindeki test videoları kullanılarak sınıflandırıcının farklı duyarlılık oranlarına göre yüz tanıma performanslarının değerlendirilmesi amacıyla uygulama yapılmıştır. Akan görüntüde aynı kişi varsa ve sahne belirlenen eşik değerinden fazla değişmiyorsa karşılaştırma oranı düşük kabul edilir. Kısacası, sahne değiştirme bilgisinden yararlanılarak uygun koşullar altında sınıflandırıcı eşik değeri değiştirilir. Dlib kitaplığına göre varsayılan tolerans değeri 0,6'dır ve daha düşük oranlar yüz karşılaştırmalarını daha katı veya hale getirir (King, 2017). Arka planda bir değişiklik olmadığı sürece gösterge yüksek hassasiyet değeri ile hareket eder. Her değişiklikte, hassasiyet oranı varsayılan değere sıfırlanır. Bu çalışmada deneyler için kullanılan donanım/işletim sistemi bilgileri Tablo 4.6'da sunulmuştur. Deneysel uygulama ayarları aşağıdaki gibidir:

$Tf_{imm} = \%30$,

varsayılan_eşik=0,6,

karşılaştırm_yeni_eşik=0,55 (yüksek hassasiyet değeri).

Tablo 4.6. Donanım ve işletim sistemi bilgi tablosu

Donanım	
Üretici ve Model	: HP-EliteBook-840-G4
Sistem Tipi	: x64-based PC
İşlemci	: Intel(R)Core(TM) i5-7300U CPU @ 2.60GHz
Grafik Kartı	: Intel® HD Graphics 620
Fiziksel Bellek	: 8.00 GB (RAM)
İşletim Sistemi	: Ubuntu 16.04 LTS 64-bit

Deneysel çalışmadan elde edilen sonuçlar aşağıdaki bazı grafik ve tablolarda sunulmuştur. İlk olarak, farklı eşik değerleriyle karşılaştırmalardan elde edilen bulgular paylaşılmıştır.

Şekil 4.25(a)'da gösterilen örneklerde sahne değişim modülü kullanılmamaktadır. Sahnede herhangi bir değişiklik olmamasına rağmen, varsayılan karşılaştırma eşiği veya duyarlılığı ile test edilerek bazı görüntü karelerinde yanlış tanımların (FP) meydana geldiği görülmüştür. Önerilen modeli değişmeyen sahnelere uygulayarak ve eşik değerini düşürerek, Şekil 4.25(b)'de görüldüğü gibi aynı karelerde doğru tanımlar elde edildiği gözlemlenmiştir.

Çerçeve: ...	2306	2311	2316	2321	2326 ...
Adı	Mesut Özil	Mesut Özil	Mali	Mesut Özil	Mesut Özil
Eşik:0.6	TP	TP	FP	TP	TP



(a)

Adı	Mesut Özil	Mesut Özil	Mesut Özil	Mesut Özil	Mesut Özil
Eşik:0.55	TP	TP	TP	TP	TP



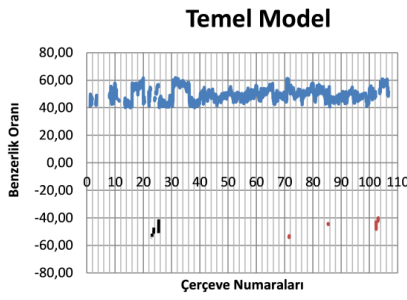
(b)

Şekil 4.25. Gerçek zamanlı video görüntülerde farklı eşik değerlerinin karşılaştırılması.

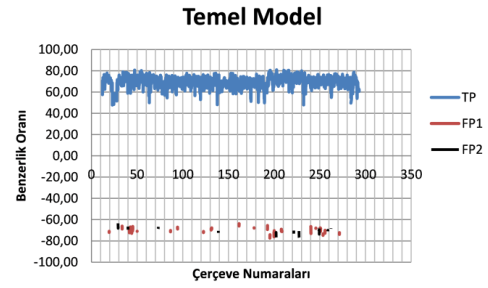
Şekil 4.26 ve Şekil 4.27'de video boyunca meydana gelen tüm TP ve FP oranları iki farklı karşılaştırma yöntemine göre grafiksel olarak gösterilmiştir. Bu bölümlerde en çok yanlış tanınan kişileri FP1 ve FP2 olarak seçerek iki yöntem karşılaştırılmıştır. Grafiklerdeki değişiklikleri daha anlaşılır kılmak için FP oranları negatif bölgede, TP oranları ise pozitif bölgede gösterilmiştir.

Şekil 4.26(a)'daki 0,6 olarak ele alınan varsayılan eşik değeri, Şekil 4.26(b)'deki değişmeyen sahne bölümlerinde dinamik olarak 0,55'e dönüştürülmüştür. Önerilen modele göre, Tablo 4.5'teki video no-3 için TP oranlarının arttığı, FP oranlarının ise azaldığı gözlemlenebilmektedir.

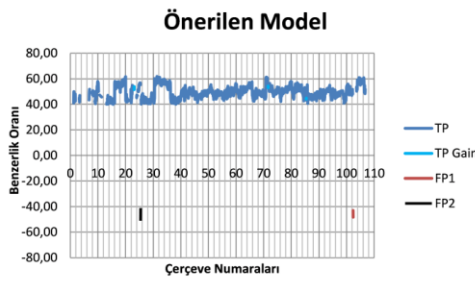
SCI yöntemiyle yapılan başka bir test videosunda Şekil 4.27(a)'da negatif bölgede gösterilen FP oranlarının Şekil 4.27(b)'de tamamen kaybolduğu görülmektedir.



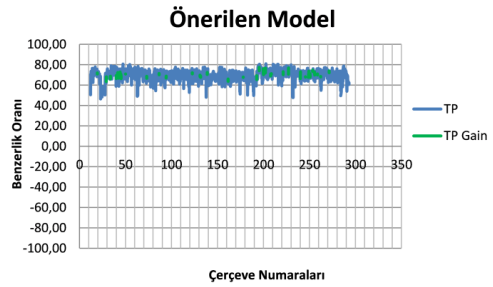
(a)



(a)



(b)



(b)

Şekil 4.26. Video No-3 için farklı model eşiklerinin karşılaştırılması.

Şekil 4.27. Video No-7 için farklı model eşiklerinin karşılaştırılması.

İkinci olarak, duyarlılık oranının farklı eşik değerlerinde nasıl değiştiğini göstermek için uygulama sonuçları Tablo 4.7'de verilmiştir. Bu tabloda, en popüler değerlendirme metrikleri kullanılarak doğruluk değerlendirmesi yapılmıştır. Bu değerlendirme metrikleri şu şekilde tanımlanır: Kesinlik (P), Hassasiyet (R) ve F1 puanı (F1) (Lee vd., 2022).

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (3)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (4)$$

$$F_1 = 2x \frac{PxR}{P + R} \quad (5)$$

Tablo 4.7. Önerilen modelin performans tablosu

Video No	Hassasiyet		Kesinlik		Doğruluk		F1-Puanı	
	Eşik (0.65)	Eşik (0.6)	Eşik (0.65)	Eşik (0.6)	Eşik (0.65)	Eşik (0.6)	Eşik (0.65)	Eşik (0.6)
1	0.9929	0.9929	0.9957	0.9957	0.9913	0.9913	0.9943	0.9943
2	0.9997	1.0000	1.0000	1.0000	0.9997	1.0000	0.9999	1.0000
3	0.9834	0.9885	0.9970	0.9978	0.9814	0.9867	0.9902	0.9931
4	0.9314	0.9279	0.9794	0.9717	0.9694	0.9572	0.9548	0.9493
5	0.8746	0.9314	0.9684	0.9855	0.9688	0.9733	0.9191	0.9576
6	0.9867	0.9808	0.9952	0.9980	0.9852	0.9897	0.9909	0.9894
7	0.9996	0.9998	1.0000	1.0000	0.9996	0.9998	0.9998	0.9999
8	0.9884	0.9927	0.9992	1.0000	0.9879	0.9927	0.9938	0.9963
9	0.9413	0.9730	0.9784	0.9930	0.9617	0.9769	0.9595	0.9829
10	0.9925	0.9924	0.9979	0.9979	0.9943	0.9928	0.9952	0.9951
11	0.9437	0.9889	0.9888	1.0000	0.9575	0.9910	0.9657	0.9944
12	0.9685	0.9796	0.9915	0.9902	0.9700	0.9754	0.9798	0.9849
Ortalama	0.9669	0.9790	0.9910	0.9942	0.9806	0.9856	0.9786	0.9864
								0.9925
								0.9925

Üçüncü olarak, modelimize benzer CFSM (Chu vd., 2022) adı verilen farklı bir yaklaşıma ait deneysel sonuçlar aşağıdaki tabloda gösterilmiştir. Bu tablo ile Tablo 4.9'daki kendi modelimizden elde edilen sonuçlar iki farklı değer ile incelenerek doğruluk karşılaştırması yapılmıştır.

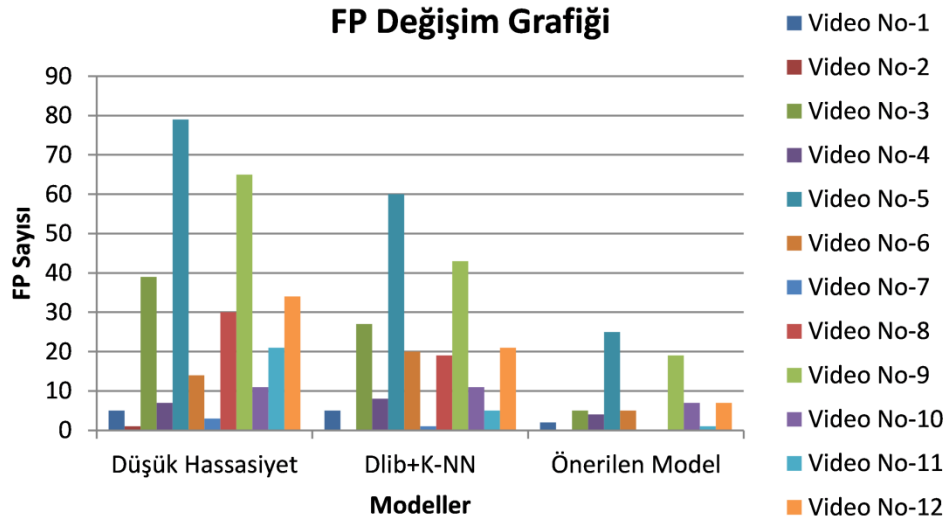
Tablo 4.8. CFSM modelinin performans tablosu

Video No	Video Adı	CFSM
		Doğruluk
1	"Changeling"	0.92
2	"Before Sunset"	0.96
3	"The International"	0.97
4	"Patriot Games"	0.95
5	"Say Anything"	0.96
6	"Instant Family"	0.86
7	"Funny People"	0.88
8	"City Slickers"	0.74
9	"Uncle drew"	0.85
10	"Animal house"	0.87

Tablo 4.9. Dlib+K-nn ve SCI modellerinin performans tablosu

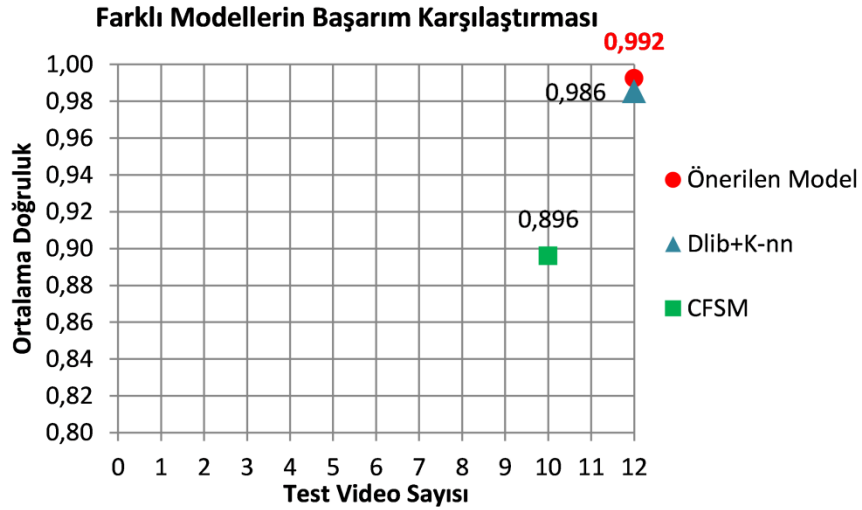
Video No	Video Adı	Dlib+K-nn	SCI
		Doğruluk	Doğruluk
1	"Adile Naşit"	0.99	1.00
2	"Aleksandra Wozniak"	1.00	1.00
3	"Ata Demirer"	0.99	0.99
4	"Maria Sharapova"	0.96	0.97
5	"Martin Palermo"	0.97	0.99
6	"Matt Smith"	0.99	0.99
7	"Mehmet Öz"-1	1.00	1.00
8	"Mehmet Öz"-2	0.99	1.00
9	"Mesut Özil"	0.98	1.00
10	"Morgan Freeman"	0.99	0.99
11	"Naim Süleymanoğlu"	0.99	1.00
12	"Wozniak ve Sharapova"	0.98	0.98

Ayrıca, tüm test videolarında farklı değerler içeren modellere göre FP değişim grafiği aşağıdaki gibidir.



Şekil 4.28. Farklı modellerin FP değişim grafiği.

Son olarak, farklı modellerin ortalama performans karşılaştırma grafiği aşağıdaki gibidir.



Şekil 4.29. Farklı modellerin ortalama performans karşılaştırması.

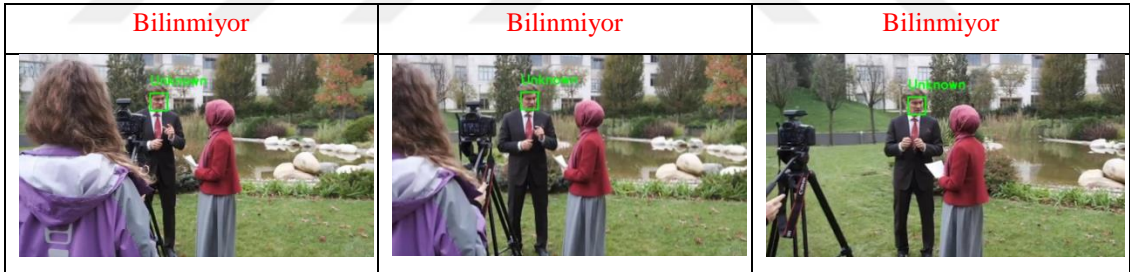
Burada dikkat edilmesi gereken bir husus da iki tane eşik değeri kullanılmasıdır. Bunlardan biri, sahnede değişikliğin olup olmadığına karar vermek için kullanılan eşik değeridir. Bu değer sahne değişikliği algılama modülünde ele alınmaktadır.

Diğeri ise, sınıflandırıcı aşamasında tespiti yapılmaya çalışılan yüz görüntüsünün veya bir diğeri ifadeyle yüz vektörünün kimliğini belirlemede kullanılan eşik değeridir.

4.4.7 Uygulamanın Analiz Edilmesi

Uygulamadan elde edilen görüntüler özellikle TP (True Positive) ve FP (False Positive) değerleri dikkate alınarak sıralanmıştır. Çünkü çalışmanın ana amacı, yanlış tanıma, yani FP değerlerinin azaltılarak başarıya olumlu etki yapmasını sağlamaktır. Önerilen yöntemin uygulanmasından önce ve sonra video görüntüler üzerinden elde edilen bazı örnekler Şekil 4.30 ile Şekil 4.31’de sırasıyla gösterilmektedir.

Aşağıdaki şekildeki hareketli kamera ile çekilen örnek sahnelerde sahne değişimi modülü kullanmadan, aşırı hassas bir oran ile yapılan karşılaştırmalarda, bazı görüntü çerçevelerinde kişinin tanınmaması durumu olduğu görülmüştür. Sahne değişimi göstergesi kullanmadan sadece statik modeldeki hassasiyet oranları gibi bir değer ile sistemin çalıştırılması, veri kümesindeki hiçbir yüz vektörü ile eşleşme yapılamadığı için bilinmeyen kişi olarak atama yapılmasına sebep olmuştur.



Şekil 4.30. Sabit, aşırı hassas oran ile elde edilen görüntüler.

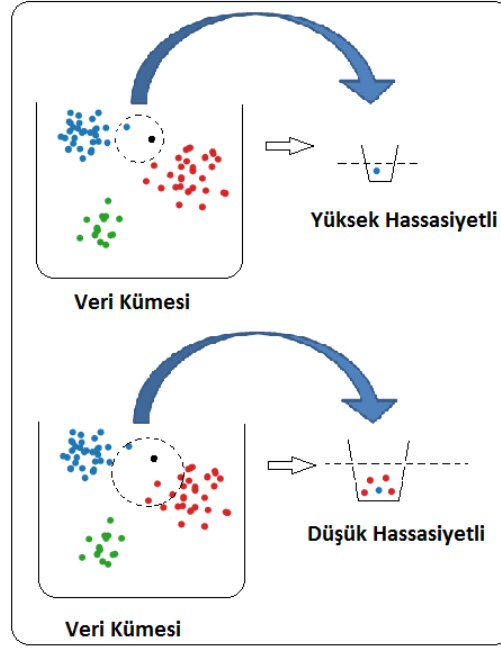
Şekil 4.31’de ise aynı video görüntüsünde sahne değişim göstergesi kullanılmıştır. Sahne değişim göstergesi ile ilk baştaki görüntülerde odak noktasından sapmadan en ideal karşılaştırma eşik değeri ile hareket ederken, sadece gösterilen son karede sahne büyük oranda değişime uğradığı için kişinin tanınmaması durumu oluşmuştur.



Şekil 4.31. Uzak çekim, değişken hassasiyet oranı ile elde edilen görüntüler.

4.4.8 Değişken Hassasiyet Oranları Analizi

Yapılan testlerden elde edilen görüntüler incelendiğinde sınıflandırıcı eşik değerinin dinamik olarak değiştirilmesinin sağladığı bazı faydalar mevcuttur. Bu durum k-en yakın komşu algoritmasının çalışma prensibinden kaynaklanmaktadır (Kutlugün, 2017). Sınıflandırılması istenen bir yüz görüntüsü vektörü, çok yüksek hassasiyet oranı ile ele alındığında, aslında bu vektöre en çok benzeyen vektörlerin veri kümesinden bulunması istenmektedir. Böyle bir vektör bazen veri kümesinde bulunabilir, ancak bazen de bu kadar çok benzeyen bir vektör bulunmadığı için tanınmayan yüz (Bilinmeyen kişi) sonucu döndürülür. Çünkü çok yüksek hassasiyet değeri ile neredeyse yüz tanıma sisteminin, bu yüz görüntüsünün aynısını bulup-getirmesi istenilmektedir. Sistem tarafından kendine en yakın olanı numune getireceği için doğru örnek yerine hatalı tanımlama yapacak bir örneği getirebilme olasılığı mevcuttur. Şekil 4.32’de, üst kısımda gösterilen görüntü buna örnek gösterilebilir. Şekle göre, çok yüksek hassasiyet seçeneği ile çok küçük boyutlu bir kova tercih edilmiş olmakta ve çok sınırlı sayıda örneği veri kümesi havuzundan çıkarmaya çalışmaktadır. Bu durumda kovanın boş olarak geriye dönmesi de olasıdır. Hassasiyet oranının daha makul oranlara düşürülmesi ile sınıflandırıcı, veri kümesindeki kendi sınıfından olan örneklerle, doğru sınıflandırılabilir için daha fazla hak tanıma şeklinde davranmaktadır. Düşük hassasiyet tercihinde kovanın boyutu büyütülmekte ve veri kümesinden kendi sınıfındaki örnekleri de içerecek şekilde daha çok örnek toplanabilme olasılığı oluşabilmektedir. Şekil 4.32’de alt kısımdaki gösterim buna örnek olarak verilebilir.



Şekil 4.32. Eşik hassasiyet örnekleme.

Tüm bunlara rağmen tanınması beklenen yüz görüntüsünün kendi sınıf sınırının dışındaki bir konumdan dolayı istenmeyen tanımlamalar oluşabilir. İyi eğitilmemiş bir model ile iç içe geçebilecek sınıflar hatalı sınıflandırmalara sebep oluşturabilir. Ayrıca, bazı durumlarda karşılaştırılacak örnek sayısının artmasına sebep olunacağı ve hesaplama maliyetinin artacağı da göz önünde bulundurulmalıdır. Bu tür modeller, veri kümesindeki numunelerin sayısına bağımlı modellerdir. Çünkü bir sınıftan çok fazla sayıda ve dağınık bir biçimde örnek bulunması, sayıca fazla olan sınıfın daha baskın hale gelerek tercih edilmesi şeklinde sonuçlanabilecektir.

BEŞİNCİ BÖLÜM

SONUÇ

Temsili öğrenme adı verilen derin metrik öğrenme algoritmaları yüz tanıma sistemlerinde sıklıkla tercih edilmektedir. Yüz tanıma sistemlerinde ilk olarak, görüntü üzerinde bulunan kişi veya kişilere ait yüz tespiti yapılmaktadır. Tespit edilen yüz görüntüleri bir metrik öğrenme modeli yardımıyla eğitilerek yüzü ifade eden vektörler kümesine dönüştürülmektedir. Yüz veri kümesinde kimliklendirilerek saklanmakta olan bu değerler tanınması istenen hedef yüz vektörleri ile belli ölçütlerde karşılaştırılarak sınıflandırma veya tanıma görevi yerine getirmektedir. Bu aşamaya kadar gerçek zamanlı sistemler için kayan pencereler üzerindeki sadece yüz bölgeleri ile işlem yapılmakta, yüz bölgeleri dışında kalan bölgeler dikkate alınmamaktadır. Önerilen bu çalışma ile birbirini takip eden sahnelerdeki değişiklikler de dikkate alınarak gerçek zamanlı yüz tanıma sistemine bir parametre şeklinde dâhil edilmektedir. Sahne değişim göstergesi parametresi ile sınıflandırıcı eşik değeri arasında bağ kurularak karşılaştırma eşik değerinin dinamik olarak değişmesi amaçlanmaktadır. Böylece, aynı kişiye ait görüntülerde poz açısı veya ışık şiddeti gibi etmenler değişse de odak noktası daha az değişime uğrayarak yanlış pozitif tanımlamaların önüne geçilebilecektir.

Sonuç olarak, sahne ortamında belirgin düzeyde değişmeler olmadıkça, yüz tanıma sistemi tarafından kayıp olarak nitelendirilebilecek hataların azaltılması sağlanmış ve daha tutarlı bir yüz tanıma sistemi elde edilmiştir. Özetle, Yanlış Pozitif olarak tespit edilen karelerin sayısı azaltılarak genel başarıma olumlu katkı sağlanmaktadır.

Önerilen yöntemden elde edilen sonuçlar, yanlış yüz tanımanın büyük ölçüde azaldığını ve genel başarımın ciddi şekilde arttığını göstermektedir.

5.1 Sınırlamalar ve Öneriler

Bu çalışma ile yapılan uygulama ve test sonuçları gerçek zamanlı yüz tanıma sistemlerine yeni bir bakış açısı kazandırarak farklı alanlarda ve amaçla kullanılan bir başka yöntemin yüz tanıma sistemlerine dâhil edilmesini sağlamıştır. Böylece yüz tanıma sisteminin daha kararlı çalışmasına olanağı doğmuştur. Ancak önerilen

sistemin bazı sakıncaları veya sınırlılıkları da mevcuttur. Bunlardan bazıları aşağıda maddeler halinde açıklanmıştır.

- Yüz tanıma sistemleri genel olarak yüz tespit sistemlerine önemli ölçüde bağlıdır. Çünkü yüz tespit sistemlerinin düzgün sonuçlar üretmemesi sistemin bütününe olumsuz etkilemektedir.
- Önerilen çalışmada, yüz tespit sisteminin yanında sahne değişim modülüne de bağımlılık oluşmaktadır. Bazı durumlarda bu durum dezavantaj oluşturabilmektedir. Örneğin; sahne değişimi olması durumunda sahnenin değişmediği sonucu ile hareket eden bir sistem, sınıflandırıcı aşamasında kullanılan karşılaştırma eşik değerini düşük veya yüksek hassasiyet ile çalıştırmaya devam edecek ve böylece doğru tanımlanabilecek iken hatalı tanımlamalara sebep olabilecektir.
- Veri kümesindeki dengesiz dağılım oluşturacak sayıda örneklerin bulunması hatalı tanımlamalara sebep olabilecektir.
- Sabit arka planlı sahne görüntüleri içeren videolarda kişiler değişip farklı yüz görüntüleri ortaya çıksa da sahne değişimi algılanmadığı için hassasiyet oranlarını düzenlemek için farklı çözümlerin geliştirilmesi gerekebilir.
- Sahne değişim göstergesi ile sadece sınıflandırıcı aşamasında kullanılan karşılaştırma eşik değerini farklı değerler ile kullanmanın dışında daha başka parametre veya özelliklerden yararlanılması çok daha verimli çalışan yüz tanıma sistemleri elde etmeyi sağlayabilir.
- Bu çalışmada sistem maliyet hesabı yapılmamıştır. Yani, sistemin çok hızlı çalışmasından ziyade hatalı bazı durumların azaltılarak genel başarımın artırılması hedeflenmiştir. Buna rağmen, standart bir donanımla bile kabul edilebilir oranlarda performans değerlerine ulaşılabilmektedir.

KAYNAKÇA

- Abu-El-Haija, vd. (2016). Youtube-8m: A large-scale video classification benchmark. *arXiv preprint*, arXiv:1609.08675.
- Akın, F. (2019). *Derin Öğrenme Tabanlı Yüz Tanıma Sisteminin Geliştirilmesi*. Elazığ: Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Allison, Robert S., Yoshitaka Fujii, ve Laurie M. Wilcox. (2020). Effects of Motion Picture Frame Rate on Material and Texture Appearance. *IEEE Transactions on Broadcasting* 6(2): 360-371.
- Amos, B., Ludwiczuk, B. ve Satyanarayanan, M. (2016). Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science, Tech. Rep.*, CMU-CS(16):118.
- Ananthakumar, A. (2018). Efficient face and gesture recognition for time sensitive application. *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*: 117-120. IEEE.
- Arpacı, S.A. ve Varlı S. (2022). Retinal Vessel Segmentation with the Mixup Data Augmentation Method. *Journal of the Presidency of Türkiye Health Institutes*, 5(1): 41-50.
- Atasoy, H. (2016). *Çok Çekirdekli İşlemcilerde Yeni Bir Paralel Görüntü İşleme Yöntemi Ve Yöntemin Yüz Tanıma Problemine Uygulanması*. Hatay: İskenderun Teknik Üniversitesi, Mühendislik ve Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Atik, M. (2020). Lokal özellik temelli yöntemler kullanılarak 3B yüz tanıma ve doğruluk analizi. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 36(1): 359-372.
- Ayata, F. (2020). *İçerik Tabanlı Görüntü Erişim Yöntemleriyle Aile Bireylerinde Yüz Tanıma Sistemi*. Van: Van Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü
- Badem, H. (2017). *Derin Öğrenme Yöntemleri Kullanarak Hiperspektral İmgelerin Sınıflandırılmasına Yönelik Yeni Yaklaşımlar*. Kayseri: Erciyes Üniversitesi,

Fen Bilimleri Enstitüsü, Doktora Tezi.

Balakrishnama, S. ve Ganapathiraju, A. (1998). Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, (18): 1-8.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends in Machine Learning*, 2(1): 1-127.

Bilgiç, A., Kurban, O.C. ve Yıldırım, T. (2017). Face recognition classifier based on dimension reduction in deep learning properties. *25th Signal Processing and Communications Applications Conference*, p:1-4. A: IEEE DOI: 10.1109/SIU.2017.7960368.

Bozik, T. (2019). *Yüz Tanıma Sistemlerinde Canlılık Analizi*. İstanbul: Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.

Buslaev, Alexander, vd. (2018). Alumentations: fast and flexible image augmentations. *Information*, 11(2):125.

Chen, X.W. ve Lin X. (2014). Big Data Deep Learning: Challenges and Perspectives. *IEEE*, (2): 514-525.

Chu SL, Chen CF, Zheng YC. (2022). CFSM: a novel frame analyzing mechanism for real-time face recognition system on the embedded system. *Multimedia Tools and Applications*, 81(2): 1867-1891.

Corrow, S.L., Dalrymple, K.A., & Barton, J.J. (2016). Prosopagnosia: current perspectives. *Eye and brain*, (8): 165.

Cuculo, Vittorio vd. (2019). Robust Single-Sample Face Recognition by Sparsity-Driven Sub-Dictionary Learning Using Deep Features. *Sensors*, 19(1): 146.

Çalık, N. (2019). *Geniş Ölçekli Veriler Üzerinde Sınıflandırma ve Bölütleme Amaçlı Evrimsel Sinir Ağı ve İstatistiksel Modellerin Geliştirilmesi*. İstanbul: Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.

Çalık, N., Kurban, O.C., Yılmaz, A.R., Ata, L.D., Yıldırım, T. (2017). Signature recognition application based on deep learning. *Signal Processing and Communications Applications Conference (SIU)*, p: 1-4. IEEE, doi: 10.1109/SIU.2017.7960454.

Dalal, N. ve Triggs, B. (2005). Histograms of oriented gradients for human detection.

- IEEE Int. Conference on Computer Vision Pattern Recognition*, p: 886–893.
inProc. IEEE Int. Conference on Computer Vision Pattern Recognition.
- Danielsson, P. (1980). Euclidean distance mapping. *Computer Graphics and image processing*, 14(3): 227-248.
- Doğan, F. ve Türkoğlu, İ. (2018). Derin Öğrenme Algoritmalarının Yaprak Sınıflandırma Başarımlarının Karşılaştırılması. *Sakarya University Journal of Computer and Information Sciences*, 1(1): 10-21.
- Doğan, F. ve Türkoğlu, İ. (2019). Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(2): 409-445.
- Duman, E. ve Akın, F. (2019). Vgg16-CNN Parametrelerinin Model Başarısına Etkilerinin Değerlendirilmesi. *Uluslararası Mardin Artuklu Multidisipliner Çalışmalar Kongresi*, p: 26-34.
- Fukushima, K. (1980). A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.*, 36(4): 193–202.
- Gandhi, A. (2021). *Data Augmentation | How to use Deep Learning when you have Limited Data—Part 2*. <https://nanonets.com/blog/tag/data-augmentation/>: <https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/> adresinden alınmıştır (Erişim tarihi: 15 Kasım 2022)
- Geitgey, A. (2017). *Face Recognition*. github.com: https://github.com/ageitgey/face_recognition#face-recognition adresinden alınmıştır (Erişim tarihi: 2 Şubat 2019)
- Goodfellow, I.J., Warde-Farley, D., Lamblin, P., Dumoulin, V., Mirza, M., Pascanu, R., Bergstra, J., Bastien, F. & Bengio, Y. (2013). Pylearn2: a machine learning research library. *arXiv preprint*, arXiv:1308.4214.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Güllü, M.K., Urhan, O. ve Ertürk, S. (2004). Dissolve detection for archive films. *Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference*, p: 510-513. Kuşadası, Aydın, Türkiye: IEEE DOI:

10.1109/SIU.2004.1338577.

- H. A. Song ve S.-Y. Lee. (2013). Hierarchical Representation Using NMF. *International Conference on Neural Information Processing*, p: 466–473.
- Hanilçi, A. (2019). *Evrşimsel sinir ağıları kullanılarak EKG ve yüz tabanlı biyometrik tanıma*. Bursa Teknik Üniversitesi, Yüksek Lisans Tezi.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, p: 770-778. IEEE.
- Hinton, G.E., & Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786): 504-507.
- Hinton, G.E., Osindero, S., & Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7): 1527-1554.
- Huang GB, vd. (2008). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. *In Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*.
- Hubel, D.H. ve Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.*, 195(1): 215–243.
- İnik, Ö. ve Ülker, E. (2017). Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Journal of Scientific Research*, 6(3): 85-104.
- Kahraman, F., Kurt, B. ve Gokmen, M. (2005). Active appearance model based face recognition. *In Proceedings of the IEEE 13th Signal Processing and Communications Applications Conference*, p: 483-486. IEEE.
- Kaplan, A. (2018). *Gerçek ve Yarı Gerçek Zamanlı Yüz Tespit Etme*. Elazığ: Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Karakaya, F.M., Altun, H. ve Çavuşlu, M. (2009). Gerçek Zamanlı Nesne Tanıma Uygulamaları için HOG Algoritmasının FPGA Tabanlı Gömülü Sistem Uyarlaması. *Signal Processing and Communications Applications Conference*. IEEE.
- Karaman, İ. (2019). *Bulut Bilişim Tabanlı Gerçek Zamanlı Yüz Tanıma Sistemi*.

Kırıkkale: Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.

Kazemi, V. ve Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, IEEE.

King, D. (2009). Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research*, (10): 1755–1758.

King, D. (2017). *High Quality Face Recognition with Deep Metric Learning*. blog.dlib.net: <http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html> adresinden alınmıştır (Erişim tarihi: 2 Şubat 2019)

Kong, P., Mancas, M., Thuon, N., Kheang, S., & Gosselin, B. (2018). Do deep-learning saliency models really model saliency? *In 2018 25th IEEE International Conference on Image Processing (ICIP)*, p: 2331-2335, IEEE.

Koptur, M. (2017). *Makine Öğrenimi*. t-SNE ile Manifold Learning: <https://makineogrenimi.wordpress.com/2017/07/09/t-sne-ile-manifold-learning/> adresinden alınmıştır (Erişim tarihi: 2 Ekim 2020)

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, (25): 1097-1105.

Kutlugün, M. (2017). *Gözetimli makine öğrenmesi yoluyla türe göre metinden ses sentezleme*. İstanbul: İstanbul Sabahattin Zaim Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.

Kutlugün, M.A. & Şirin, Y. (2018b). A novel approach improvement framework for text to speech synthesis. *In 2018 26th Signal Processing and Communications Applications Conference (SIU)*, p: 1-4. IEEE. DOI:10.1109/SIU.2018.8404828.

Kutlugün, M.A. & Şirin, Y. (2020). Anlamli ve Benzer Olmayan Türkçe Metinler Üretmek için N-Gram Yöntemi ile İstatistiksel ve Kural Tabanlı Yaklaşımın Birlikte Kullanımı. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 22(65): 331-342. DOI: 10.21205/deufmd.2020226504.

Kutlugün, M.A. ve Şirin, Y. (2018a). Turkish meaningful text generation with class

- based n-gram model. *26th Signal Processing and Communications Applications Conference (SIU)*, p: 1-4. IEEE. DOI: 10.1109/SIU.2018.8404801.
- Kutlugün, M.A. ve Şirin, Y. (2023). Augmenting the training database with the method of gradual similarity ratios in the face recognition systems. *Digital Signal Processing*, <https://doi.org/10.1016/j.dsp.2023.103967>.
- Kutlugün, M.A. ve Şirin, Y. (2017). Dimension reduction selection for increasing naturalness in speech synthesis. *25th Signal Processing and Communications Applications Conference (SIU)*, p: 1-4. IEEE. DOI: 10.1109/SIU.2017.7960573.
- Kutlugün, M.A., Şirin, Y. ve Karakaya, M.A. (2019). The Effects of Augmented Training Dataset on Performance of Convolutional Neural Networks in Face Recognition System. *Federated Conference on Computer Science and Information Systems (FedCSIS)*, p: 929-932. Leipzig, Germany: IEEE. DOI: 10.15439/2019F181.
- Lawrence, S., Giles, C.L., Tsoi, A.C. ve Back, A.D. (1997). Face recognition: A convolutional neural-network approach. *8(1)*: 98-113. *IEEE transactions on neural networks*.
- LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10).
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86(11)*: 2278-2324.
- Lee, H.S. ve Cho, S.I. (2022). Luminance Level of Histogram-Based Scene-Change Detection for Frame Rate Up-Conversion. *IEEE Access*, (10): 15968-15977.
- Leonard J. Press, O.D., Faaq, Fccovd. (2014). *Prosopagnosia and Google Glass*. The Visionhelp Blog: <https://visionhelp.wordpress.com/2014/01/19/prosopagnosia-and-google-glass/> adresinden alınmıştır (Erişim Tarihi: 23 Aralık 2021)
- Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. *In Proceedings of the IEEE Conference*

- on Computer Vision and Pattern Recognition*, p: 5325-5334.
- Liu, Si vd. (2015). Matching-cnn meets knn: Quasi-parametric human parsing. *Proceedings of the IEEE conference on computer vision and pattern recognition*, p: 1419-1427.
- Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, p: 91-110.
- Lu, X. (2003). Image analysis for face recognition. *Personal notes*, 36.
- Lv, J.J., Shao, X.H., Huang, J.S., Zhou, X.D., & Zhou, X. (2017). Data augmentation for face recognition. *Neurocomputing*, (230): 184-196.
- Moghaddam, B.V. (1999). Bayesian image retrieval in biometric databases. *International Conference on Multimedia Computing and Systems*, (2): 610-615. IEEE.
- Moor, B.C. (1981). Principle component analysis in linear system. *IEEE Trans Autom Control*, (11): 17-32.
- Olson, D.L. ve Delen, D. (2008). *Advanced data mining techniques*. Springer Science & Business Media.
- Özkaya N. ve Sağiroğlu Ş. (2008). Face recognition from fingerprints. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 23(4): 785-793.
- Özcan, H. (2014). *Çok düşük çözünürlüklü yüz imgelerinde derin öğrenme uygulamaları*. İstanbul: Deniz Harp Okulu Komutanlığı, Deniz Bilimleri ve Mühendisliği Enstitüsü, Yüksek Lisans Tezi.
- Öznel, G. (2019). *Yüz Analizine Dayalı Derin Öğrenme Tabanlı Bir İlgi Tespit Sisteminin Gerçekleştirilmesi*. Sakarya: Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Öznel, İ. (2018). *Kısmi ve Tam Yüz Görüntüleri Üzerinde Makine Öğrenmesi Yöntemleriyle Yüz İfadesi Tespiti*. Sakarya: Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Peemen, M. (2018). *Convolutional Neural Network (CNN)*. NVIDIA Developer: <https://developer.nvidia.com/discover/convolutional-neural-network> adresinden alınmıştır (Erişim Tarihi: 7 Ocak 2019)

- Ranzato, M.A., Boureau, Y.L., & LeCun, Y. (2007). Sparse Feature Learning for Deep Belief Networks. *Proc. Adv. Neural Inf. Process. Syst.*, (20): 1185-1192.
- Rosebrock, A. (2018). *Pyimagesearch*. Face recognition with OpenCV, Python, and deep learning: <https://www.pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/> adresinden alınmıştır (Erişim Tarihi: 11 Temmuz 2019)
- Sağlam, G. (2008). *Hareketli görüntülerde gerçek zaman yüz tanıma yaklaşımları*. Zonguldak: Zonguldak Karaelmas Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Salman, M. (2018). *Hiperspektral ve Lidar Verilerinin Öznitelik ve Karar Seviyelerinde Tümlleştirilmesi ve Derin Evrişimli Sinir Ağlarıyla Sınıflandırılması*. Ankara: Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Sangdoon Y., Dongyoon H., Seong J.O., Sanghyuk C., Junsuk C. ve Youngjoon Y. (2019). Cutmix: Regularization strategy to train strong classifiers with localizable features. *In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, p: 6023-6032. IEEE.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p: 815–823.
- Serengil, Ş. (2021). *DeepFace – The Most Popular Open Source Facial Recognition Library*. viso.ai: <https://viso.ai/computer-vision/deepface/> adresinden alınmıştır (Erişim Tarihi: 26 Mart 2022)
- Sharma, A. (2018). *Introduction to Deep Learning in Python*. datacamp: <https://www.datacamp.com/tutorial/introduction-deep-learning> adresinden alınmıştır (Erişim Tarihi: 8 Haziran 2020)
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint*, arXiv:1409.1556.
- Soltanpour, S., Boufama, B. ve Wu, Q.J. (2017). A survey of local feature methods for 3D face recognition. *Pattern Recognition*, (72): 391-406.

- Söylemez, Ö. (2021). *Yüz İfadesi Tanıma Başarımı İçin İlgili Bölge Belirleme ve Çekişmeli Üretici Ağlar Kullanılarak Örnek Artırım Yöntemlerinin Geliştirilmesi*. Elazığ: Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, (15): 1929–1958.
- Subaşı, N. (2011). *Gerçek Zamanlı İnsan Yüzü Belirleme*. Edirne: Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Sufiyan, A. (2018). *Derin öğrenme tekniklerini kullanarak gerçek zamanlı yüz tanıma sisteminin geliştirilmesi*. Gaziantep: Gaziantep Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Sütçüler, E. (2006). *Gerçek zamanlı video görüntülerinden yüz bulma ve tanıma sistemi*. İstanbul: Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Şahin, F. (2020). *Simgesel Noktaları Kullanarak Şekil ve Gri Düzeyde Görünümlere Dayalı Yüksek Ayrımcı Yüz Tanıma Yöntemi*. İstanbul: İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, Doktora Tezi.
- Şahin, Ö. (2017). *iOS Platformunda Görme Engelliler İçin TL Tanıma Uygulaması*. Konya: Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Şeker, A., Diri, B. ve Balık, H.H. (2017). Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme. *Gazi Mühendislik Bilimleri Dergisi*, 3(3): 47-64.
- Tang, Y. (2013). Deep learning using linear support vector machines. *arXiv preprint*, arXiv:1306.0239.
- Uçar, M. (2019). *Gerçek Zamanlı Görüntü İşleme ile Sınıf Ortamındaki Öğrencilerin Tanınması ve Dikkat Dağılımlıklarının Tespiti*. İskenderun, Hatay: İskenderun Teknik Üniversitesi, Mühendislik ve Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(11).

- Viola, P.J. (2004). Robust real-time object detection. *International Journal of Computer Vision*, (57): 137–154.
- Vu HN, Nguyen MH, Pham C. (2021). Masked face recognition with convolutional neural networks and local binary patterns. *Applied Intelligence*, p: 1-16.
- Watson, A.B. (2013). High frame rates and human vision: A view through the window of visibility. *SMPTE Motion Imaging Journal*, 122(2): 18–32.
- Xiao, H.R. (2017). Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint*, arXiv: 1708.07747.
- Yaniv, T., Ming, Y., Marc'Aurelio, R., Lior, W. (2014). DeepFace: Closing the gap to human-level performance in face verification. *In Computer Vision and Pattern Recognition (CVPR)*, p: 1701–1708. IEEE.
- Yazıcı, İ. (2008). *Gerçek zamanlı kiosk yüz tanıma sistemi*. Kocaeli: Gebze Yüksek Teknoloji Enstitüsü, Mühendislik ve Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.
- Yıldırım, M. (2019). *Derin Öğrenme Teknikleri Kullanarak Yüz Tanıma Tabanlı Müşteri Doğrulama ile Bankamatiklerde Sahtekârlık Tespiti*. Konya: Konya Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi.
- Zhang, H., Cisse, M., Dauphin, Y.N. ve Lopez-Paz, D. (2017). mixup: Beyond empirical risk minimization. *arXiv preprint*, arXiv: 1710.09412.
- Zheng, X., Chalasani, T., Ghosal, K., Lutz, S., & Smolic, A. (2019). Stada: Style transfer as data augmentation. *arXiv preprint*, arXiv:1909.01056.

ÖZGEÇMİŞ

Mehmet Ali KUTLUGÜN

A. EĞİTİM

Mehmet Ali Kutlugün 1980 yılında İstanbul'da doğdu. İlk ve orta öğrenimini İstanbul'da tamamladı. 2000 yılında Süleyman Demirel Üniversitesi Bilgisayar Donanımı ön lisans programından, 2008 yılında Anadolu Üniversitesi İşletme lisans ve 2015 yılında İstanbul Sabahattin Zaim Üniversitesi Bilgisayar Mühendisliği lisans programlarından mezun oldu. 2018 yılında İstanbul Sabahattin Zaim Üniversitesi Bilgisayar Bilimi ve Mühendisliği yüksek lisans programından mezun olduktan sonra aynı üniversitede doktora programına başladı.

B. MESLEKİ DENEYİM

Net Bilgisayar Yazılım Hizmetleri, BEKO Elektronik A.Ş. ve Erenköy Ruh ve Sinir Hastalıkları Hastanesi gibi firmaların bilgi-işlem birimlerinde çalıştı. Son olarak, 2012 yılından beri Türk Hava Yolları A.O. firmasının bilgi işlem biriminde çalışmaya devam etmektedir.

C. YAYINLAR

Kutlugün M.A. ve Şirin Y. (2017). *Dimension reduction selection for increasing naturalness in speech synthesis*. 25th Signal Processing and Communications Applications Conference (SIU 2017), Antalya, IEEE. DOI:10.1109/SIU.2017.7960573

Kutlugün M.A. ve Çakır M.Y. (2017). *Yapay Sinir Ağları ve K-En Yakın Komşu Algoritmalarının Birlikte Çalışma Tekniği (Ensemble) ile Metin Türü Tanıma*. İstanbul: XXII. Türkiye'de İnternet Konferansı (inet'tr17), Bahçeşehir Üniversitesi.

- Kutlugün M.A. ve Çakır M.Y. (2017). *Derin Sinirsel Ağlarla Konuşma Tespiti ve Cinsiyet Tahmini*, İstanbul: XXII. Türkiye’de İnternet Konferansı (inet’tr17), Bahçeşehir Üniversitesi.
- Kutlugün, M. A. (2017). *Gözetimli makine öğrenmesi yoluyla türe göre metinden ses sentezleme*. İstanbul Sabahattin Zaim Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı. Yüksek Lisans Tezi. URI: <https://hdl.handle.net/20.500.12436/569>
- Kutlugün, M. A., & Şirin, Y. (2018). *A novel approach improvement framework for text to speech synthesis*. In 2018 26th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE. DOI:10.1109/SIU.2018.8404828
- Kutlugün, M. A., & Şirin, Y. (2018). *Turkish meaningful text generation with class based n-gram model*. In 2018 26th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE. DOI:10.1109/SIU.2018.8404801
- Kutlugün, M. A., Şirin, Y., & Karakaya, M. (2019). *The effects of augmented training dataset on performance of convolutional neural networks in face recognition system*. In 2019 Federated Conference on Computer Science and Information Systems (FedCSIS) (pp. 929-932). IEEE. DOI:10.15439/2019F181
- Kutlugün, M. A., & Şirin, Y. (2020). *Anlamli ve Benzer Olmayan Türkçe Metinler Üretmek için N-Gram Yöntemi ile İstatistiksel ve Kural Tabanlı Yaklaşımın Birlikte Kullanımı*. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 22(65), 331-342. DOI: 10.21205/deufmd.2020226504
- Kutlugün, M. A., & Şirin, Y. (2023). *Augmenting the Training Database with the Method of Gradual Similarity Ratios in the Face Recognition Systems*. Digital Signal Processing, 103967. DOI:10.1016/j.dsp.2023.103967
- Kutlugün, M. A., & Şirin, Y. (2023). *Reducing false positive rate with the help of scene change indicator in deep learning based real-time face recognition systems*. Multimedia Tools and Applications.