

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

AÇIKLANABİLİR YAPAY ZEKA İLE BENIGN
PAROKSİSMAL POZİSYONEL VERTİGO TANISINDA
KARAR DESTEK SİSTEMİNİN GELİŞTİRİLMESİ

DOKTORA TEZİ

Hasibe Büşra AYTEKİN

İstanbul
Ekim-2025

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

AÇIKLANABİLİR YAPAY ZEKA İLE BENIGN PAROKSİSMAL
POZİSYONEL VERTİGO TANISINDA KARAR DESTEK
SİSTEMİNİN GELİŞTİRİLMESİ

DOKTORA TEZİ

Hasibe Büşra AYTEKİN

Tez Danışmanı

Prof. Dr. Tarık ÖZKUL

İkinci Tez Danışmanı

Dr. Öğr. Üyesi Aydın Tarık ZENGİN

İstanbul

Ekim-2025

Lisansüstü Eğitim Enstitüsü Müdürlüğüne,

Bu çalışma, jürimiz tarafından Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Bilimleri ve Mühendisliği (%30 İngilizce) Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman Prof. Dr. Tarık ÖZKUL

Üye Doç. Dr. Mohammed VADİ

Üye Dr. Öğr. Üyesi Şengül BAYRAK

Üye Dr. Öğr. Üyesi Muhammed DAVUD

Üye Doç. Dr. Ülviye HACIZADE

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Erhan İÇENER
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİMİ

Doktora tezi olarak hazırladığım “**Açıklanabilir Yapay Zeka ile Benign Paroksizmal Pozisyonel Vertigo Tanısında Karar Destek Sisteminin Geliştirilmesi**” adlı çalışmanın öneri aşamasından sonuçlandığı aşamaya kadar geçen süreçte bilimsel etiğe ve akademik kurallara özenle uyduğumu, tez içindeki tüm bilgileri bilimsel ahlak ve gelenek çerçevesinde elde ettiğimi, tez yazım kurallarına uygun olarak hazırladığımı, bu çalışmamda doğrudan veya dolaylı olarak yaptığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu beyan ederim.

Hasibe Büşra AYTEKİN

ÖN SÖZ

Tez çalışmam süresince bilgi ve deneyimleriyle bana yol gösteren değerli hocalarım Prof. Dr. Tarık Özkul ve Dr. Aydın Tarık Zengin'e; klinik bilgileriyle beni yönlendiren ve tezimin bilimsel temellerinin güçlenmesine katkı sağlayan Prof. Dr. Murat Haluk Özkul'a ve bu süreçte desteğini esirgemeyen tüm hocalarıma içtenlikle teşekkür ederim.

Akademik hayatım boyunca sevgisi, fedakârlığı ve dualarıyla yanımda olan kıymetli aileme; her koşulda sabrı, anlayışı ve desteğiyle bana güç veren sevgili eşime; doktora yolculuğum boyunca yanımda olduğunu hissettiren değerli arkadaşlarıma gönülden teşekkür ederim.

Bu çalışmanın, alanına katkı sağlayacak nitelikte bir adım olmasını temenni ederim.

Hasibe Büşra AYTEKİN

İstanbul, 2025

ÖZET

AÇIKLANABİLİR YAPAY ZEKA İLE BENIGN PAROKSİSMAL POZİSYONEL VERTİGO TANISINDA KARAR DESTEK SİSTEMİNİN GELİŞTİRİLMESİ

Hasibe Büşra AYTEKİN

Doktora, Bilgisayar Bilimleri ve Mühendisliği (%30 İngilizce)

Tez Danışmanı: Prof. Dr. Tarık ÖZKUL

İkinci Tez Danışmanı: Dr. Öğr. Üyesi Aydın Tarık ZENGİN

Ekim-2025, 100 +XIII Sayfa

Benign Paroksizmal Pozisyonel Vertigo (BPPV), belirli baş hareketleriyle ortaya çıkan kısa süreli baş dönmesi ataklarıyla kendini gösteren yaygın bir periferik vestibüler bozukluktur. Erken tanı, hastaların yaşam kalitesini artırmak ve tedavi sürecini belirlemek açısından kritik önemdedir. Tanı süreci, özellikle nistagmusa ilişkin istemsiz ve ritmik göz hareketlerinin dikkatli değerlendirilmesine bağlıdır. Ancak, göz kapakları veya kirpiklerden kaynaklanan gölgeler gibi çevresel etkiler mevcut tanı yöntemlerinin etkinliğini azaltabilmektedir. Bu çalışma, BPPV tanısında doğruluğu artırmak için en etkili segmentasyon yöntemini araştırmıştır. Bu amaçla YOLOv8, YOLOv11 ve YOLOv12 modelleri kullanılmış; ek olarak Segment Anything Model 2 (SAM2) de karşılaştırmaya dahil edilmiştir. Hareket örüntülerini incelemek için BPPV vakalarına ait video verisine optik akış analizi uygulanmış, elde edilen veriler CNN-GRU hibrit modeliyle dört nistagmus türüne ayrılarak sınıflandırılmıştır. Hekimin karar sürecini desteklemek amacıyla, modele Yavaş Faz Hızı (Slow Phase Velocity, SPV) değerlerine dayalı bulanık mantık yapısı entegre edilmiştir. Analizler, SAM2 tabanlı segmentasyon yaklaşımının %94 doğruluk ile en yüksek başarıyı sağladığını ve sistemin hem teknik doğruluk hem de açıklanabilirlik açısından güçlü bir potansiyel sunduğunu göstermiştir.

Anahtar Kelimeler: BPPV, SAM2, Optik Akış, Nistagmus Sınıflandırma, Açıklanabilir Yapay Zeka, CNN-GRU, Bulanık Mantık.

ABSTRACT

DEVELOPING A DECISION SUPPORT SYSTEM FOR BENIGN PAROXYSMAL POSITIONAL VERTIGO DIAGNOSIS WITH EXPLAINABLE ARTIFICIAL INTELLIGENCE

Hasibe Büşra AYTEKİN

PhD, Computer Science and Engineering (%30 English)

Supervisor: Prof. Dr. Tarık ÖZKUL

Cosupervisor: Asst. Prof. Dr. Aydın Tarık ZENGİN

October-2025, 100 +XIII Pages

Benign Paroxysmal Positional Vertigo (BPPV) is a common peripheral vestibular disorder characterized by short episodes of dizziness triggered by specific head movements. Early diagnosis is critical for improving patients' quality of life and guiding appropriate treatment. The diagnostic process largely depends on the careful evaluation of involuntary and rhythmic eye movements associated with nystagmus. However, environmental factors such as eyelid or eyelash shadows may reduce the effectiveness of current diagnostic methods. This study aimed to identify the most effective segmentation method to improve diagnostic accuracy in BPPV. For this purpose, YOLOv8, YOLOv11, and YOLOv12 models were employed, and the Segment Anything Model 2 (SAM2) was also included for comparison. To analyze motion patterns, optical flow was applied to video data of BPPV cases, and the extracted features were classified into four types of nystagmus using a CNN-GRU hybrid model. To support clinical decision-making, a fuzzy logic structure based on Slow Phase Velocity (SPV) values was integrated into the model. The analyses revealed that the SAM2-based segmentation approach achieved the highest accuracy with 94%, demonstrating strong potential in terms of both technical reliability and explainability.

Keywords: BPPV, SAM2, Optical Flow, Nystagmus Classification, Explainable Artificial Intelligence, CNN-GRU, Fuzzy Logic.

İÇİNDEKİLER

TEZ ONAYI	i
BİLİMSEL ETİK BİLDİRİMİ.....	ii
ÖN SÖZ	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER	vi
TABLolar LİSTESİ.....	ix
ŞEKİLLER LİSTESİ.....	x
KISALTMALAR	xii
BİRİNCİ BÖLÜM.....	1
GİRİŞ	1
İKİNCİ BÖLÜM	6
BENIGN PAROKSİSMAL POZİSYONEL VERTİGO	6
2.1 BPPV'nin Tanımı.....	6
2.2 Klinik Tanı Süreçleri.....	7
ÜÇÜNCÜ BÖLÜM	9
LİTERATÜR TARAMASI	9
DÖRDÜNCÜ BÖLÜM	13
MATERYAL VE YÖNTEM.....	13
4.1. YOLO Segmentasyon Modelleri	13
4.1.1. YOLOv8 Segmentasyon Modeli	14

4.1.2. YOLOv11 Segmentasyon Modeli	15
4.1.3. YOLOv12 Segmentasyon Modeli	16
4.1.4. Ground Truth Çıkarma	17
4.1.5. YOLO Model Eğitimi.....	19
4.2. Segment Anything Model 2	20
4.3. Optik Akış Yöntemi	22
4.4. Sınıflandırma.....	26
4.4.1. Convolutional Neural Networks	26
4.4.2. Gated Recurrent Unit (GRU).....	30
4.4.3. CNN-GRU Hibrit Model	31
4.4.4. Sınıflandırma Başarı Ölçütleri.....	32
4.5. Açıklanabilir Yapay Zeka	35
4.5.1. Yavaş Faz Hızı (Slow Phase Velocity - SPV).....	37
4.5.2. Bulanık Mantık	39
BEŞİNCİ BÖLÜM	44
DENEYSEL SONUÇLAR VE ANALİZLER	44
5.1. Veri seti ve Ön işleme	44
5.2. YOLO Segmentasyon Yöntemlerinin Karşılaştırılması	47
5.3. YOLO ve SAM2 Segmentasyon Performansının Değerlendirilmesi	52
5.4. Optik Akış Analizi ve Hareket Örüntüleri	61
5.5. CNN-GRU Sınıflandırma Performansı	65
5.6. Bulanık Mantık Tabanlı Yorumlama ve Karar Desteği	73
SONUÇ.....	81

KAYNAKÇA	83
EKLER.....	97
ÖZGEÇMİŞ.....	99



TABLolar LİSTESİ

Tablo 5.1: BPPV veri kümesinde her bir nistagmus sınıfına ait örnek sayısı.....	46
Tablo 5.2: YOLO segmentasyon modellerine ait performans metrikleri.	48
Tablo 5.3: CASIA-Iris-Interval veri kümesini kullanan çalışmalar arasında segmentasyon performansı karşılaştırması; CAD: Circular Active Discs, CHT: Circular Hough Transform, ID: Integro-Differential Operator, F1: F1 Skoru.....	51
Tablo 5.4: Farklı segmentasyon yöntemleri kullanılarak elde edilen CNN-GRU modelinin sınıflandırma performansı.	66
Tablo 5.5: YOLOv8-seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.	70
Tablo 5.6: YOLOv11-Seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.	70
Tablo 5.7: YOLOv12-Seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.	71
Tablo 5.8: SAM2 segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.	71
Tablo 5.9: BPPV'de nistagmus sınıflandırması için önerilen yöntemin derin öğrenme tabanlı yaklaşımlarla karşılaştırılması.....	72

ŞEKİLLER LİSTESİ

Şekil 4.1: CASIA-Iris-Interval veri kümesine ait örnek göz bebeği görüntüleri.....	17
Şekil 4.2: Gözün anatomik yapısını gösteren ve segmentasyon işleminde hedeflenen iris (dış çember) ile pupil (iç çember) bölgeleri.....	18
Şekil 4.3: SAM2 uygulanan veri kümesinden örnek video kareleri.....	21
Şekil 4.4: Ardışık hareket eden bir topun yer değiştirme vektörünü gösteren ok.	22
Şekil 4.5: Lukas-Kanade yöntemi uygulanmış görseller.....	24
Şekil 4.6: Sıralı video kareleri ve yoğun optik akışları.	25
Şekil 4.7: HSV uzayındaki renk kodları.....	25
Şekil 4.8: Örnek CNN yapısı.....	27
Şekil 4.9: Tam Bağlantılı Katman Örneği.....	30
Şekil 4.10: Önerilen CNN-GRU Hibrit Model Mimarisi.....	31
Şekil 4.11: Görüntü segmentasyonunda tahmin ve gerçek maskelerin karşılaştırılması. Model tarafından tahmin edilen segmentasyon maskesi (a), Gerçek (ground-truth) segmentasyon maskesi (b), Tahmin ve gerçek maskelerin üst üste bindirilmiş hali (c).	34
Şekil 4.12: Bulanık mantık hareket bazlı üyelik fonksiyonları.....	41
Şekil 4.13: Çıktı üyelik fonksiyonları ve fuzzy skor kategorilendirmesi (0-10).....	42
Şekil 4.14: Üyelik derecesine göre defüzyon örneği.....	42
Şekil 5.1: Veri kümesinden ham görüntü karesi: (a) kırılmış göz bölgesi: (b).....	45
Şekil 5.2: YOLO modellerinin 50 epoch boyunca CASIA-Iris-Interval veri kümesi üzerindeki segmentasyon kayıp (segmentation loss) grafiği.....	49
Şekil 5.3: YOLO modellerinin 50 epoch boyunca CASIA-Iris-Interval veri kümesi üzerindeki maske mAP (50–95) skor grafiği.....	49
Şekil 5.4: YOLO modellerinin doğruluk ve gecikme karşılaştırması.....	50
Şekil 5.5: Şiddetli nistagmus hareket senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.....	53
Şekil 5.6: Düşük göz kapağı ve sınırlı görüş senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.....	55
Şekil 5.7: Kısmen kapalı göz kapağı ve şiddetli nistagmus senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.....	56
Şekil 5.8: Yoğun kirpik ve sınırlı göz açıklığı senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.....	58

Şekil 5.9: Makyaj kaynaklı zorluk senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.....	59
Şekil 5.10: Şiddetli nistagmus hareket senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi. ...	62
Şekil 5.11: Yoğun kirpik ve sınırlı göz açıklığı senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.	63
Şekil 5.12: Düşük göz kapağı ve sınırlı görüş senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.	63
Şekil 5.13: Kısmen kapalı göz kapağı ve şiddetli nistagmus senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.	64
Şekil 5.14: Makyaj kaynaklı zorluk senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi. ...	64
Şekil 5.15: Farklı segmentasyon yöntemleriyle (YOLOv8-Seg, YOLOv11-Seg, YOLOv12-Seg ve SAM2) eğitilen CNN-GRU modeline ait eğitim (train) ve doğrulama (validation-val) doğruluk eğrileri.....	66
Şekil 5.16: Farklı segmentasyon yöntemleriyle (YOLOv8-Seg, YOLOv11-Seg, YOLOv12-Seg ve SAM2) eğitilen CNN-GRU modeline ait eğitim (train) ve doğrulama (validation-val) kayıp (loss) eğrileri.	67
Şekil 5.17: YOLOv8-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.	68
Şekil 5.18: YOLOv11-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.	68
Şekil 5.19: YOLOv12-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.	69
Şekil 5.20: SAM2 tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.....	69
Şekil 5.21: Horizontal nistagmus içeren video kaydına ait grafik analizleri.	74
Şekil 5.22: Vertical nistagmus içeren video kaydına ait grafik analizleri.	75
Şekil 5.23: Horizontal video analiz çıktısının son karesine ait ekran görüntüsü.....	78
Şekil 5.24: Vertical video analiz çıktısının son karesine ait ekran görüntüsü.	78
Şekil 5.25: Torsiyonel video analiz çıktısının son karesine ait ekran görüntüsü.....	79

KISALTMALAR

ASC: Anterior Semicircular Canal

AUROC: Area Under the Receiver Operating Characteristic Curve

BiLSTM: Bidirectional Long Short-Term Memory

BPPV: Benign Paroksismal Pozisyonel Vertigo

CAD: Circular Active Discs

CASIA: Chinese Academy of Sciences' Institute of Automation

CCTU: Clockwise-Toward Utricle

CHT: Circular Hough Transform

CNN: Convolutional Neural Networks (Konvolüsyonel Sinir Ağları)

COCO: Common Objects in Context

CTU: Counterclockwise-Toward Utricle

FC: Fully Connected

FCN: Fully Convolutional Network

FLOPs: Floating Point Operations per Second

FN: False Negative

FP: False Positive

FPS: Frame Per Second

GPU: Graphics Processing Unit

GRU: Gated Recurrent Unit

HSV: Hue, Saturation, Value

HT: Hough Transform

ICPR: International Conference on Pattern Recognition

ID: Integro-Differential Operator

IoU: Intersection over Union

mAP: mean Average Precision

NIR: Near Infrared (Yakın Kızılötesi)

PC: Posterior Canal

RGB: Red, Green, Blue

RMS: Robotik Manevra Sistemi

RNN: Recurrent Neural Network

SAM: Segment Anything Model

SAM 2: Segment Anything Model 2

SC: Superior Canal

SVM: Support Vector Machine

TN: True Negative

TP: True Positive

t.y.: Tarih Yok

Val: Validation

VNG: Videonystagmography

YOLO: You Only Look Once

BİRİNCİ BÖLÜM

GİRİŞ

Vertigo, iç kulak denge organlarında ve bu organların merkezi bağlantılarında meydana gelen bozukluklardan kaynaklanan bir hareket hissidir. Genellikle iyi huylu olmasına rağmen bazen ciddi etkileri olabilir. Vertigo olan bir kişinin, hastalığın altında yatan denge organı bozukluğu doğru tanı ile etkili bir şekilde tedavi edilmesi mümkündür. Vertigo, periferik ve merkezi olarak sınıflandırılır. Merkezi vertigo, merkezi sinir sisteminin denge merkezlerini etkileyen bozukluklardan kaynaklanır. Periferik vertigo ise iç kulak ya da orta kulaktaki denge merkezinin etkilenmesinden kaynaklanan denge bozukluğudur (Kabade vd., 2021).

Periferik vertigonun en yaygın nedenlerinden biri BPPV'dir. BPPV, yaygın bir halk sağlığı sorunu olarak kabul edilmektedir (Yetiser ve Ince, 2015). Semptomlar, başın pozisyonundaki değişiklikler üzerine tetiklenen dönme hissi olarak tanımlanmıştır (Bhattacharyya vd., 2017). BPPV'nin baş dönmesi semptomu, iç kulağın yarım daire biçimli kanalları olarak adlandırılan yerde kalsiyum birikintileri (kristal) ile açıklanmaktadır. Şiddetli baş dönmesi semptomu kısa sürelidir ve gözün istemsiz olarak belirli yönlere hareket ettiği (nistagmus) bir durumdur (Kabade vd., 2021). Kristalin bulunduğu kanala özgü olarak göz hareketleri farklılık göstermektedir. Gözlerin anormal hareketleri, doktorun BPPV bozukluğunun türünü belirlemede yardımcı olur.

BPPV, belirli baş hareketleriyle tetiklenen kısa süreli baş dönmesi ataklarıyla karakterize edilen en yaygın vestibüler bozukluklardan biridir (Ozkul, 2022). Kişisel yaşam kalitesini olumsuz etkileyen bu yaygın durum, aynı zamanda kamu sağlık sistemleri üzerinde de önemli bir yük oluşturmaktadır. BPPV, iç kulaktaki yarım daire kanallarına yer değiştirmiş otokoniaların (kalsiyum kristalleri) girmesiyle oluşur; bu da tekrarlayan vertigo ataklarına ve günlük aktivitelerde işlevsel kısıtlılıklara neden olur (Ozkul, 2022).

Gözlerin herhangi bir anormal hareketine bağlı olarak, otoloji uzmanları her vakayı belirli bir tipe göre kategorize edebilir ve uygun tedaviyi uygular. Ancak gevşek kristaller 6 vestibüler kanaldan herhangi birinde olabilir ve her kanal farklı tipte bir nistagmusa neden olur. Bu nedenle nistagmus hareketleri oldukça komplike olabilir ve

tanısı oldukça zor olabilir. Bu yüzden doğru tanı konulana kadar çoğu hasta gereksiz testlere tabi tutulur (Öztürk vd., 2019).

BPPV'nin klinik önemi, yalnızca bireysel semptomların yönetimiyle sınırlı değildir; aynı zamanda bu durumun yanlış tanı veya tedavi edilmemesi, düşmeler ve buna bağlı travmalar gibi ikincil sağlık sorunlarına da yol açabilmektedir. Özellikle yaşlı bireylerde sık görülmesi, toplum sağlığı açısından BPPV'nin doğru tanı ve tedavisini daha da önemli kılmaktadır. Bu bağlamda, tanı sürecinin hassasiyeti, hem bireylerin yaşam kalitesini artırmak hem de sağlık sistemindeki yükü azaltmak adına kritik bir rol oynamaktadır.

BPPV tipi vertigo, nüfusun %8'ini etkiler ve her yıl nüfusun %1,6'sı BPPV atakları için acil servise başvurmaktadır. BPPV'nin acil serviste yanlış tanı oranı %74-81 olarak bildirilmektedir. ABD'de her yıl 5,5 milyon hasta acil servise başvuruyor ve BPPV'nin sağlık sistemine maliyeti yılda 4,4 milyar dolar olarak belirtilmektedir. İstatistiklere göre BPPV hastaları hastaneye 7 kez geliyor ve iyileşmeleri ortalama 70 ay sürüyor. BPPV, yaşam kalitesini önemli ölçüde düşüren hastalıklardan biridir. Nistagmus hareketleri oldukça komplike ve tanısı oldukça zor olabilir. Bu, acil serviste bildirilen %74-81 yanlış tanı oranı ile kanıtlanmıştır (Ozkul, 2022).

BPPV'nin doğru şekilde tanı edilmesi, etkili bir tedavi sürecinin yürütülmesi açısından kritik öneme sahiptir. Bu tanı sürecinin temel unsurlarından biri ise, ritmik ve istemsiz göz hareketleri olan nistagmusun değerlendirilmesidir (Fife vd., 2008). Nistagmusun yönü, süresi ve örüntüsü, BPPV'nin alt tiplerinin ayırt edilmesinde belirleyici rol oynamaktadır. Ancak bu değerlendirme, düşük göz kapakları, gölgelenmeler ve ışık yansımaları gibi görsel engellerle zorlaşabilir (Milman, Yellin ve Aperstein, 2025). Ayrıca, nistagmusun ani ve hızlı doğası, hareket bulanıklığı ve çevresel gürültüye karşı hassas olan geleneksel görüntü işleme yöntemlerinin güvenilir analiz sağlamasını zorlaştırmaktadır.

Klinikte nistagmus değerlendirmesi çoğunlukla uzman hekimin deneyimine dayanır. Ancak bu süreç hem zaman alabilir hem de göz hareketlerinin hızlı ve düzensiz doğası nedeniyle hataya açık hale gelebilir. Bu nedenle, göz hareketlerini otomatik ve nesnel şekilde değerlendirebilen sistemlerin geliştirilmesi, tanı süreçlerini daha güvenilir hale getirebilir.

Gelişen yapay zekâ teknolojileri, trafik izleme, tarım, güvenlik ve otonom sistemler gibi birçok farklı alanda yaygın biçimde uygulanmaktadır (Azimjonov, Özmen & Varan, 2023; Dogru & Zegin, 2022; Mohanty vd., 2016; Wang & Deng, 2021;). Bu teknolojilerin bir alt dalı olan bilgisayarla görü teknikleri ise, hastalıkların tanı ve tedavisinde önemli ilerlemeler sağlamaktadır. Büyük klinik veri kümelerinde manuel olarak model çıkarmak zor ve yavaş bir süreçtir. Daha hızlı ve daha doğru bir değerlendirme oluşturmak için verileri görselleştirmek, anlamak ve sınıflandırmak için bilgisayarlı görü teknikleri aracı olarak kullanılabilir.

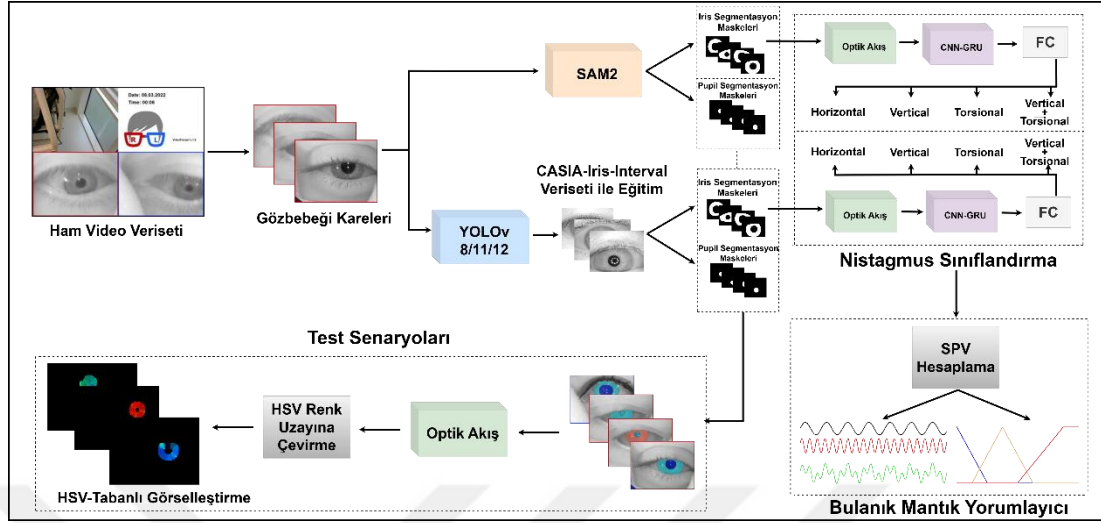
Yapay zekâ ve bilgisayarla görü alanındaki son gelişmeler, tıbbi görüntüleme de görsel tanı süreçlerini önemli ölçüde iyileştirmiştir. Özellikle derin öğrenme modelleri, görsel verilerin yorumlanmasına dayalı görevlerde yüksek doğruluk oranları elde etmiştir (Kabade vd., 2021; Öztürk vd., 2019). Ancak BPPV gibi hareket dinamiklerinin merkezi olduğu durumlarda, durağan görüntülere dayalı yaklaşımlar genellikle yetersiz kalmaktadır. Hareket yönü ve zamansal örüntüleri birlikte ele alan yöntemler, tanı doğruluğunu artırılmasında gereklidir.

Bu çalışmanın amacı, manevralar yapılarak nistagmus bilgisi kaydedilen hastaların, geliştirdiğimiz derin öğrenme yöntemi ile BPPV bozukluğu tanısı için maliyeti azaltmak ve hastanın doğru tanı ile iyileşme hızını artırmaktır.

Bu çerçevede, klinik uygulamalarda video-okülografi araçları kullanılmaya başlanmış ve derin öğrenme sistemlerinin düşük kaliteli video görüntülerinde dahi otomatik nistagmus tespiti konusunda potansiyel taşıdığı gösterilmiştir (Newman-Toker vd., 2013; Wagle vd., 2022). Görüntü işleme sistemlerinin başarısı yalnızca kullandıkları algoritmalara değil, aynı zamanda gerçek dünyada karşılaşılan koşullara ne kadar uyum sağlayabildiklerine de bağlıdır. Özellikle göz bölgesinde yapılan analizlerde, düşük görüntü kalitesi veya çevresel engeller gibi etkenler segmentasyon doğruluğunu olumsuz etkileyebilir. Kısmen kapalı göz kapakları, yoğun kirpikler ve nistagmusa bağlı hızlı göz hareketleri gibi durumlar manuel tespiti zorlaştıran yaygın engeller arasında yer almaktadır.

Bu zorlukları simüle etmek amacıyla, literatürde bildirilen yaygın segmentasyon hatalarına ve klinik uzmanlığın öngörülerine dayalı olarak beş gerçek dünya senaryosu tasarlanmıştır. Her bir senaryo, görsel olarak belirsiz veya hata yapmaya açık bir

durumu temsil etmekte ve çevresel faktörlerin segmentasyon doğruluğu üzerindeki etkisini değerlendirmek üzere yapılandırılmıştır.



Şekil 1.1: BPPV tanısında göz hareketi analizine yönelik önerilen yöntemin genel akışı.

Zorlu klinik senaryolar altında kaydedilen ham video verileri, göz karelerinin çıkarılması amacıyla işlenmiştir. İki farklı segmentasyon stratejisi uygulanmıştır: (i) SAM2 modeli ile doğrudan göz karesi segmentasyonu ve (ii) CASIA-Iris-Interval veri kümesi ile eğitilmiş YOLOv8/11/12 modelleri kullanılarak iris segmentasyonu. Segmentasyon çıktıları, senaryoya özgü testlerle değerlendirilmiştir. Ardından, segmentlenen göz bebeği bölgelerine optik akış analizi uygulanmış ve BPPV ile ilişkili nistagmus hareketleri, HSV tabanlı renk kodlaması ile görselleştirilerek karakterize edilmiştir. Son aşamada, optik akış tensörlerinden elde edilen öznitelik haritaları, CNN-GRU modeli ve ardından gelen tam bağlantılı (FC) katman aracılığıyla dört farklı nistagmus yönüne sınıflandırılmıştır. Şekil 1-1’de BPPV tanısında göz hareketi analizine yönelik önerilen iş akışının genel görünümü sunulmaktadır.

Bu çalışmada, zorlu iris tespiti senaryolarında derin öğrenmeye dayalı çeşitli segmentasyon modelleri karşılaştırılarak etkililikleri değerlendirilmiştir. YOLOv8, YOLOv11 ve YOLOv12 modelleri, açık erişimli bir iris veri kümesi üzerinde eğitilmiş ve hız, doğruluk ve model karmaşıklığı açısından değerlendirilmiştir. Bu modeller, BPPV’ye özgü klinik koşullar altında SAM2 modeliyle birlikte test edilerek segmentasyon performansları analiz edilmiştir. Segmentasyon işlemi takiben, hareket örüntülerini çıkarmak amacıyla optik akış analizi uygulanmış ve elde edilen

veriler, CNN-GRU modeli kullanılarak dört farklı nistagmus türüne sınıflandırılmıştır. Bu çerçeve, segmentasyon kalitesinin hem hareket analizine hem de sınıflandırma doğruluğuna etkisini kapsamlı bir şekilde değerlendirmeye olanak sağlamıştır.

Çalışmanın temel katkısı, BPPV vakalarında tanı doğruluğunu artırmak amacıyla etkili segmentasyon yaklaşımını belirlemek için yapılan karşılaştırmalı model değerlendirmesidir. Ayrıca, çevresel koşullar ve hızlı göz hareketlerinin oluşturduğu zorlukların üstesinden gelmek için optik akış tabanlı hareket yönü tahminini entegre eden sağlam bir analiz çerçevesi sunulmuştur. Çalışmada ayrıca, geleneksel ve derin öğrenmeye dayalı yöntemlere ilişkin bir literatür taraması, önerilen yöntem ve veri kümesinin detayları, deneysel sonuçlar ile birlikte klinik yorumlar ve gelecekteki araştırma yönelimleri de yer almaktadır.

BPPV, toplumda sık karşılaşılan bir denge bozukluğu olup, tanı süreci büyük ölçüde klinisyenin gözlemlerine ve deneyimine dayanmaktadır. Bu durum, özellikle erken evrelerde ya da ayırt edilmesi güç nistagmus türlerinde tanı hatalarına yol açabilmektedir. Bu tez kapsamında geliştirilen görüntü tabanlı analiz yöntemi, nistagmus hareketlerinin değerlendirilmesini mümkün kılarak tanı sürecine destek olmayı amaçlamaktadır. Derin öğrenme tabanlı segmentasyon ve sınıflandırma modellerinin birlikte kullanılmasıyla, hem görüntü işleme hem de yapay zekâ alanında anlamlı bir katkı sağlanmakta; ayrıca bu yöntemlerin klinik karar destek sistemlerine entegrasyonuna yönelik önemli bir altyapı ortaya konmaktadır.

Bu tez toplam beş bölümden oluşmaktadır. Birinci bölümde, çalışmanın problemi, amacı ve kapsamı açıklanarak araştırmanın dayandığı temel gerekçeler sunulmuştur. İkinci bölümde, BPPV, nistagmus hareketleri ve bu alanda geliştirilen yapay zekâ temelli yöntemler üzerine güncel literatür incelenmiştir. Üçüncü bölümde, çalışmada kullanılan YOLO segmentasyon modelleri, SAM2, optik akış yöntemi ve CNN-GRU hibrit sınıflandırma modeli detaylı olarak açıklanmış; veri hazırlama ve model eğitim süreçleri sunulmuştur. Dördüncü bölümde, geliştirilen yöntemlerin çeşitli metriklerle değerlendirildiği deneysel çalışmalar yer almakta olup, segmentasyon ve sınıflandırma performansları analiz edilmiştir. Beşinci ve son bölümde ise elde edilen bulgular özetlenmiş, çalışmanın katkıları tartışılmış ve gelecekte yapılabilecek araştırmalara yönelik öneriler sunulmuştur.

İKİNCİ BÖLÜM

BENIGN PAROKSİSMAL POZİSYONEL VERTİGO

2.1 BPPV'nin Tanımı

BPPV, erişkin popülasyonda en yaygın görülen periferik vestibüler hastalıktır ve tekrarlayıcı, kısa süreli vertigo ataklarıyla karakterizedir. Genellikle birkaç saniye süren bu vertigo atakları başın yerçekimi ile ilişkili hareketleriyle tetiklenir; örneğin aşağı eğilme, yukarı bakma, yataкта dönme veya yataktan kalkma gibi durumlarda ortaya çıkar. BPPV atakları özellikle yaşlı bireylerde sabahları yataktan kalkarken düşme ile sonuçlanabilecek denge bozukluklarına yol açabilir (Kim ve Zee, 2014; Van Dam vd., 2021). Bu durum, vestibüler sistemin mekanik bir bozukluğu olarak değerlendirilir ve temel olarak otolit organlarından ayrılmış olan kalsiyum karbonat kristallerinin (otokoniaların) yarım daire kanalları içine kaçması ile ilişkilidir. BPPV'nin tanısı klinik değerlendirme ile konulabilmekte, hastaların büyük çoğunluğu uygun pozisyonlama manevraları ile tedavi edilebilmektedir (Halmágyi vd., 2023).

BPPV'nin altında yatan patofizyolojik mekanizma, otokoniaların utriküler otolit membranından ayrılarak yarım daire kanallarına yerleşmesidir. En yaygın görülen form olan kanalitiaziste, bu partiküller kanal içerisinde serbest olarak hareket ederken; kupulolitik durumunda ise kupulaya yapışmış olarak bulunurlar (Kao, Parnes ve Chole, 2017). Başın pozisyon değiştirmesiyle birlikte yerçekiminin etkisiyle otokonialar kanal içinde hareket eder, bu da kanal afferentlerinin bazal aktivitesini değiştirerek anormal vestibüler sinyallerin ortaya çıkmasına yol açar. Bu anormal sinyaller, beynin vestibüler çekirdekleri tarafından yanlış yorumlanarak vertigo ve göz hareketlerinde nistagmus şeklinde yansıma bulur. Nistagmusun yönü, etkilenen kanalın düzlemiyle ilişkilidir ve rotasyon eksenini uyarılan kanalın düzlemine ortogonal olarak gözlenir (Aw vd., 2005). Bu fizyopatolojik süreçler, baş hareketi ile tetiklenen pozisyonel vertigo ve tipik göz bulguları ile klinikte tanı konmasına olanak sağlar.

BPPV'nin altında yatan patofizyolojik mekanizma, otokoniaların utriküler otolit membranından ayrılarak yarım daire kanallarına yerleşmesidir. En yaygın görülen form olan kanalitiaziste, bu partiküller kanal içerisinde serbest olarak hareket ederken; kupulolitik durumunda ise kupulaya yapışmış olarak bulunurlar (Kao, Parnes ve Chole, 2017). Başın pozisyon değiştirmesiyle birlikte yerçekiminin etkisiyle otokonialar kanal içinde hareket eder, bu da kanal afferentlerinin bazal aktivitesini

değiştirerek anormal vestibüler sinyallerin ortaya çıkmasına yol açar. Bu anormal sinyaller, beynin vestibüler çekirdekleri tarafından yanlış yorumlanarak vertigo ve göz hareketlerinde nistagmus şeklinde yansıma bulur. Nistagmusun yönü, etkilenen kanalın düzlemiyle ilişkilidir ve rotasyon eksenini uyarılan kanalın düzlemine ortogonal olarak gözlenir (Aw vd., 2005). Bu fizyopatolojik süreçler, baş hareketi ile tetiklenen pozisyonel vertigo ve tipik göz bulguları ile klinikte tanının konmasına olanak sağlar.

2.2 Klinik Tanı Süreçleri

BPPV tanısında en önemli adım, hastanın ayrıntılı semptom öyküsünün dikkatle değerlendirilmesi ve ardından yapılan muayenede doğru pozisyonel testlerin uygulanmasıdır. Bu testler sayesinde, vertigonun hangi yarım daire kanalından kaynaklandığı belirlenebilir. En sık kullanılan tanı testi Dix-Hallpike manevrasıdır ve özellikle posterior kanal BPPV'si için altın standart kabul edilir (Talmud vd., 2017).

Test sırasında hasta, muayene masasının kenarına oturur pozisyondayken, başı yaklaşık 45 derece sağa ya da sola çevrilir ve ardından hızla sırtüstü yatırılarak başı yaklaşık 20 derece sarkıtılır. Bu pozisyon değişikliğiyle birkaç saniyelik gecikmenin ardından genellikle yukarı yönlü ve torsiyonel özellikte bir nistagmus gözlenir. Nistagmusun yönü, etkilenen kanalın hangisi olduğuna dair önemli ipuçları sunar. Dix-Hallpike testi aynı zamanda anterior kanal tutulumlarında da kullanılabilir, ancak anterior kanal BPPV klinikte oldukça nadir görülmektedir.

Supine Roll testi (ya da yatış-yuvarlanma testi) ise horizontal kanal BPPV'sini değerlendirmek amacıyla uygulanır (Hızal vd., 2022). Hasta sırtüstü pozisyonda yatırıldıktan sonra, başı sırasıyla sağa ve sola 90 derece döndürülerek her iki kanalın da uyarılması sağlanır. Eğer test sırasında yatay düzlemde nistagmus ortaya çıkarsa, bu durum horizontal kanalın etkilendiğini gösterir. Nistagmusun geotropik (yer yönelimli) veya apogeotropik (yer karşıtı) olması, kristallerin kanal içindeki konumuna ve tutulum tipine göre değerlendirilir.

Daha nadir görülen anterior kanal BPPV'sinin tanısında ise straight head-hanging testi kullanılabilir (Anagnostou, Kouzi ve Spengos, 2015). Bu testte hasta oturur pozisyondayken, başı hızla geriye doğru 60 derece sarkıtılarak yatırılır. Eğer aşağıya doğru yönelen torsiyonel nistagmus izlenirse, anterior kanal tutulumundan şüphe edilir. Her üç test de klinik olarak oldukça değerlidir ve doğru uygulandığında, hem BPPV tanısını destekler hem de hangi kanalın etkilendiğini belirleyerek uygun tedavi

manevrasının seçilmesine olanak tanır. Özellikle göz hareketlerinin video-mercek sistemleriyle izlenmesi, nistagmusun yönü ve süresi gibi kritik parametrelerin daha net değerlendirilmesini sağlar ve tanı doğruluğunu artırır.

BPPV tanısında göz hareketlerinin analizi son derece önemlidir. Baş pozisyonlarının değiştirilmesiyle ortaya çıkan pozisyonel nistagmus, hem tanı koymada hem de hangi kanalın etkilendiğini belirlemede temel rol oynar (Halmágyi vd., 2023). Nistagmusun yönü, süresi, latens süresi ve yavaş faz hızı gibi parametreler, tanı için ayırt edici özellikler taşır. Örneğin posterior kanal BPPV’inde gözlenen geotropik torsiyonel ve yukarı yönlü nistagmus, tipiktir. Bu nistagmusun genellikle 2–3 saniyelik bir latens süresi sonrası ortaya çıkması yavaş faz hızına sahip olması tanı koymada yol göstericidir (Aw vd., 2005). Yatay kanal BPPV’inde ise yatış pozisyonunda her iki tarafa dönerken oluşan geotropik veya apogeotropik yatay nistagmuslar gözlenebilir. Bu göz hareketlerinin dikkatli incelenmesi ve vücut pozisyonuyla ilişkili olarak yön değiştirmesi, hastalığın lokalizasyonu ve tedavi stratejisinin belirlenmesi açısından kritiktir.

BPPV’nin klinik alt tipleri, etkilediği yarım daire kanalına göre sınıflandırılır. En sık rastlanan form posterior kanal BPPV olup, tüm BPPV olgularının yaklaşık %90’ını oluşturur (Halmágyi vd., 2023). Bu formda Dix-Hallpike testi ile tipik olarak yukarı yönlü ve torsiyonel nistagmus elde edilir. Yatay kanal BPPV, tüm olguların yaklaşık %10’unu oluşturur ve genellikle supin roll testi ile tanı konulabilir. Bu formda nistagmus daha çok yatay düzlemde, geotropik veya apogeotropik yönde olabilir. Yatay kanal tutulumunda nistagmus süresi biraz daha uzundur ve genellikle latens süresi daha kısadır (Lechner vd., 2014).

Nadir görülen bir alt tip ise anterior kanal BPPV’dir ve genellikle başı geriye düşürme pozisyonlarında aşağı yönlü nistagmus ile karakterizedir. Bunun dışında atipik BPPV formlarında kupulolitiyazis veya posterior kanalın distal kısmında bulunan otokonialar, klasik testlerde alışılmış nistagmus paternleri dışında bulgulara yol açabilir (Vannucchi vd., 2015). Bu olguların tanı ve tedavisi daha karmaşık olup, deneyimli bir nörootolog değerlendirmesi gerektirebilir.

ÜÇÜNCÜ BÖLÜM

LİTERATÜR TARAMASI

Göz bebeği tespiti ve nistagmus analizi, çeşitli metodolojik çerçeveler kullanılarak kapsamlı biçimde araştırılmıştır. Bu alandaki teknikler, geleneksel görüntü işleme yöntemlerinden modern derin öğrenme modellerine ve doğrudan video tabanlı yaklaşımlara kadar uzanmakta olup, her biri nistagmus sınıflandırmasında farklı avantajlar ve sınırlılıklar sunmaktadır. Özellikle BPPV tanısında, çevresel etkileşimleri en aza indirirken iris takibi doğruluğunu artırmaya yönelik yöntemlere olan ilgi giderek artmaktadır. Bu doğrultuda, literatür üç ana başlık altında incelenmiştir: (1) geleneksel görüntü işleme yöntemleri, (2) iris segmentasyonu ve nistagmus analizi için derin öğrenmeye dayalı teknikler ve (3) doğrudan video verisi kullanılarak yapılan hareket tabanlı yaklaşımlar.

Geleneksel görüntü işleme teknikleri, nistagmusun tespiti ve analizi için yaygın şekilde kullanılmaya devam etmektedir. Bu yöntemler genellikle video dizilerinde iris hareketini izlemek için kural tabanlı algoritmalara ve matematiksel modellere dayanır. Yaygın yaklaşımlar arasında Hough Dönüşümü (HT) ve uyarlanabilir şablon eşleme yer almakta olup, her ikisi de iris segmentasyonu ve hareket kestiriminde sıkça uygulanmaktadır (Hashemi vd., 2016; Ro, 1972). Örneğin, Slama ve ark., videonistagmografi (VNG) sisteminde göz bebeği segmentasyonu için HT yöntemini kullanmış ve yüksek tespit doğruluğu elde etmiştir; ancak hızlı göz hareketlerinde izleme performansı düşmüştür. Slama vd. (2018) ayrıca bir kişinin vestibüler bozukluğu olup olmadığını tespit etmek için nistagmusu analiz etmek üzere çok katmanlı bir sinir ağı kullanmış ve HT yöntemi ile göz hareketini sınıflandırmada başarılı sonuçlar elde etmişlerdir.

Slama vd. (2018) Hough Dönüşümü temelli yaklaşımlarına benzer şekilde, Dogru vd. (2024), BPPV hastalarında nistagmus bileşenlerini analiz etmek için hibrit bir yöntem sunmuştur. Çalışmada, göz bebeği tespiti için YOLOv8 kullanılmış; ardından yatay ve dikey hareketler polar koordinat dönüşümüyle, torsiyonel hareketler ise geleneksel şablon eşleme yöntemiyle analiz edilmiştir. Böylece derin öğrenme temelli segmentasyon ile klasik görüntü işleme tabanlı hareket analizinin birlikte kullanıldığı, uygulama odaklı bir sistem ortaya konmuştur.

Lim vd. (2019), BPPV tanısında 3B göz bebeği takibi için kütle merkezi tahmini ile elips oturtmayı birleştiren bir karar destek sistemi önermiştir. Yazarlar, hastaların on farklı pozisyondaki göz hareketlerini kaydederek oluşturdukları geniş veri setinde, ön işlem aşamasında göz bebeği merkezini ağırlık merkezine dayalı bir algoritma ile takip etmiş, Hough Dönüşümü ve elips yerleştirme algoritmalarını kullanarak göz bebeği segmentasyonu gerçekleştirmiştir. Yatay ve dikey hareketlerin yönü geçici hızın toplamı ile belirlenmiş, burulma hareketleri ise çapraz korelasyon ile ölçülmüştür. AUROC değerleri sırasıyla 0.966, 0.952 ve 0.853 olarak rapor edilmiştir. Benzer şekilde, Zhang vd. (2021), torsiyonel hareketi tespit etmek amacıyla HT ve optik akışı derin öğrenmeyle entegre etmiş; ancak düşük çözünürlüklü ve hızlı hareket içeren video verilerinde yetersiz kalmıştır.

Göz torsiyonunun güvenilir şekilde ölçümüne odaklanan bir çalışmada, iris şekli sayısal olarak modellenmiş ve bozulmalar düzeltilerek şablon eşleme yöntemiyle torsiyon açısı hesaplanmıştır. Bu yöntem, gerçekleştirilen testlerde düşük hata oranlarıyla başarılı sonuçlar vermiş ve klinik uygulamalarda kullanılabilirliğini ortaya koymuştur (Jin vd., 2020).

Bir diğer çalışmada, gözün torsiyonel hareketlerini ölçmek için yeni bir yöntem önermektedir. Yöntemde, gözün irisindeki zamansal olarak sabit kalan koyu alanlar, "Maksimum Stabil Hacimler" (Maximally Stable Volumes) adı verilen görüntü özellikleriyle takip edilir. Bu özellikler, ışık değişimlerine ve göz ile kamera pozisyonundaki değişimlere karşı dayanıklıdır. Yöntem, ışık yansımaları veya göz kapağı gibi kısmi engellemeler altında dahi güvenilir sonuçlar verirken, makak maymunlarının neredeyse özelliksiz görünen irislerinde bile başarılı şekilde uygulanabilmektedir (Ong ve Haslwanter, 2010).

Gözün torsiyonel hareketlerini yani göz küresinin bakış hattı etrafındaki dönmesini ölçmek için yeni bir yöntem tanıtılmıştır. Araştırmacılar, iris tanıma ve modern görüntü işleme tekniklerini kullanarak, invaziv olmayan, hassas, düşük gürültülü ve gerçek zamanlı çalışan bir sistem geliştirmiştir. Bu sistem, özellikle baş pozisyonuna göre nesne yönelimi algısı gibi üç boyutlu görsel algıyı anlamak ve bazı nörolojik ya da vestibüler hastalıkların tanısı açısından önemlidir. Yeni yöntem, mevcut standart olan skleral halka yöntemiyle karşılaştırıldığında daha doğru sonuçlar vermekte ve geniş bilimsel ve klinik uygulama potansiyeline sahiptir (Otero-Millan vd., 2015).

Gelişmiş derin öğrenme modelleri, göz tespiti ve hareket analizi için giderek daha fazla kullanılmakta olup, oküler hareketlerin daha hassas takibini ve nistagmus türlerinin doğru sınıflandırılmasını sağlamaktadır. Bu yöntemler, iris veya göz bebeğini hızlı ve güvenilir bir şekilde segmentlemeyi hedeflerken, çevresel gürültüye karşı duyarlılığı da azaltmayı amaçlamaktadır.

Kong vd. (2023), optik akışı segmentasyon teknikleriyle entegre ederek torsiyonel hareketin doğru tespitine odaklanan bir nistagmus sınıflandırma sistemi geliştirmiştir. Bu çalışmada iki farklı veri seti kullanılmıştır: dört sınıflı veri setinde saat yönü veya tersine burulma ile yükselen nistagmus; iki sınıflı veri setinde ise torsiyonel ve non-torsiyonel nistagmus yer almıştır. Optik akış HSV formatında çıkarılmış, her pikselin yönü ve büyüklüğü analiz edilmiştir. ConvNext-B ve ResNet101 gibi derin ağlar kullanılarak %97.75'e varan doğruluklara ulaşılmıştır.

Lu vd. (2024), göz hareketi ile baş pozisyon verisini birleştiren çok modlu bir model geliştirmiş ve bu birleşimin BPPV tanısını iyileştirdiğini göstermiştir. Çalışmada göz bebeği segmentasyonu için önceden eğitilmiş bir model kullanılmış, burulma hareketleri faz korelasyonu yöntemiyle ölçülmüştür. Iris desenine log-polar dönüşüm uygulanmış ve histogram eşitleme ile görsel detaylar artırılmıştır. Elde edilen üç ayrı modelin doğrulukları yatay %93, dikey %90.5, torsiyonel %87.95 olarak kaydedilmiştir.

Pham vd. (2022), yatay nistagmusu izleyen ve başlangıç-bitiş anlarını yüksek doğrulukla tespit eden hibrit bir derin öğrenme sistemi önermiştir. Çalışmada yedi farklı duruşta video verisi toplanmış ve toplam altı sınıf tanımlanmıştır. Veriler öncelikle ikiye ayrılarak arka kanal ayrımı yapılmış, ardından GRU tabanlı model ile sınıflandırılmıştır. Özellikle yanal kanallarda “güçlü vuruş” sinyalini yakalayamayan klasik CNN-RNN yapılarına alternatif olarak vuruş hızı ölçümüne dayalı bir algoritma geliştirilmiştir. Göz koordinatları kullanılarak vuruş hızları hesaplanmış ve SVM ile sınıflandırılmıştır. Altı sınıflı veri setinde ortalama %91 doğruluk elde edilmiştir.

Ayrıca, Poplin vd. (2018), Esteva vd. (2017) ve Carreira vd. (2017) gibi öncü çalışmalar, derin öğrenmenin tıbbi görüntülerde sınıflandırma ve ölçüm süreçlerinde etkili olduğunu vurgulamaktadır. Bu çalışmaların ışığında, BPPV gibi vestibüler bozuklukların analizinde derin öğrenme tabanlı sistemlerin umut vadettiği görülmektedir.

Bazı çalışmalar ise doğrudan video verisine dayalı nistagmus analizini, açık biçimde iris veya göz bebeği tespiti yapmadan gerçekleştirmiştir. Li ve Yang (2023), genişletilmiş ve ayırık evrişimleri BiLSTM-GRU yapısıyla birleştiren bir derin öğrenme çerçevesi ile dikey nistagmus sınıflandırması yapmıştır. Modelleri, geleneksel göz bebeği takibini atlamasına rağmen yüksek doğruluk elde etmiştir.

Wagle vd. (2022) ise, video kareleri arasındaki göz pozisyonu değişimlerini analiz ederek yatay, dikey ve torsiyonel nistagmusu tespit eden bir sistem geliştirmiştir. Modelleri, farklı koşullar altında test edilmiş ve standart göz takibi tekniklerine ihtiyaç duymadan güçlü sınıflandırma performansı göstermiştir.

Benzer şekilde, Tang ve Li (2017), BPPV tanısı için çeşitli tanı ve yeniden konumlandırma manevraları uygulamış ve video temelli analizlerle tedavi başarısını artırmayı amaçlamıştır. Ayrıca, doğrudan video karelerinden çıkarılan göz bebeği koordinatları ile hız hesaplamaları yapılmakta ve bu özellikler, GRU veya SVM gibi sınıflayıcılara aktarılmaktadır.

Özetle, göz tespiti ve nistagmus sınıflandırması için en uygun yaklaşım, büyük ölçüde uygulama bağlamına ve veri özelliklerine bağlıdır. Geleneksel yöntemler hızlı ve düşük maliyetlidir; ancak zorlu koşullarda genellikle hassasiyetten yoksundur. Derin öğrenmeye dayalı teknikler daha yüksek doğruluk ve uyarlanabilirlik sunar; yine de gerçek dünya uygulamaları için hem sağlam hem de kaynak verimli çözümler gereklidir. Bu çerçevede, tez kapsamında segmentasyonun sınıflandırma performansı üzerindeki etkisi analiz edilmiş ve elde edilen bulgular, nistagmus hareketlerinin doğru biçimde belirlenmesine yönelik geliştirilen yaklaşımların bir parçası olarak değerlendirilmiştir.

DÖRDÜNCÜ BÖLÜM

MATERYAL VE YÖNTEM

Bu çalışmada, BPPV tanısında göz segmentasyon doğruluğunu artırmak ve farklı segmentasyon yaklaşımlarının performansını değerlendirmek amacıyla hem YOLO tabanlı derin öğrenme modelleri hem de SAM2 modeli kullanılmıştır. Çalışmanın temel hedefi, çevresel faktörlerin ve hızlı nistagmus hareketlerinin segmentasyon kalitesi üzerindeki etkisini inceleyerek en uygun yöntemi belirlemektir. Segmentasyonun ardından hareket örüntülerini daha ayrıntılı analiz edebilmek için optik akış tekniği uygulanmış ve elde edilen hareket vektörleri, görselleştirme amacıyla HSV renk uzayına dönüştürülmüştür. Ayrıca, BPPV'ye özgü bir video veri kümesine uygulanan optik akış analizinden elde edilen öznitelik haritaları kullanılarak dört sınıflı nistagmus sınıflandırması gerçekleştirilmiştir. Bu yaklaşım, BPPV ile ilişkili karakteristik hareket örüntülerini belirlemeyi ve tanı sürecinde klinik karar verme süreçlerini desteklemeyi amaçlamaktadır.

4.1. YOLO Segmentasyon Modelleri

YOLO ailesi, nesne tespiti ve segmentasyon görevleri için en hızlı ve en verimli derin öğrenme mimarilerinden biri olarak kabul edilmektedir (Redmon vd., 2016). Gerçek zamanlı çalışabilme yeteneği, yüksek doğruluk oranı, düşük hesaplama maliyeti ve tek aşamalı mimarisi sayesinde, görüntü analizi uygulamalarında yaygın şekilde kullanılmaktadır. Göz hareketlerinin doğru şekilde tespiti, BPPV tanısının kritik bir bileşenidir. Ancak geleneksel görüntü işleme teknikleri, hızlı göz hareketleri veya bulaşma, düşük göz kapakları ya da yansımalar gibi çevresel bozulmalar karşısında çoğu zaman yetersiz kalmaktadır (Zhang vd., 2021).

Geleneksel YOLO modelleri büyük ölçüde sınırlayıcı kutu (bounding box) tespitine dayanmakta olup, piksel düzeyinde segmentasyon gerektiren tıbbi görüntüleme uygulamaları için yetersiz kalabilir. Bu sorunun üstesinden gelmek için, bu çalışmada sınırlayıcı kutular yerine piksel hassasiyetinde nesne maskeleri sağlayan YOLO-seg varyantları kullanılmıştır (Terven, Córdova-Esparza ve Romero-González, 2023). Bu yetenek, özellikle göz bebeği veya iris gibi yumuşak kenarlı ve çevresel etkilere duyarlı yapılar için büyük avantaj sağlamaktadır. Bu doğrultuda, bu çalışmada pupil ve iris segmentasyonu için en uygun mimariyi belirlemek amacıyla YOLOv8-seg,

YOLOv11-seg ve YOLOv12-seg olmak üzere üç segmentasyon tabanlı YOLO modeli değerlendirilmiştir. Bu modeller, mimari farklılıkları ve güncel nesne tespiti araştırmalarında en iyi performans gösteren YOLO sürümleri olmaları nedeniyle seçilmiştir.

4.1.1. YOLOv8 Segmentasyon Modeli

Ocak 2023'te Ultralytics, YOLO serisinin gelişmiş bir sürümü olan YOLOv8'i, beş farklı seçenekle kullanıcıların hizmetine sunarak bilgisayarla görme alanında dikkat çekici bir yenilik gerçekleştirdi (Sarda, Dixit ve Bhan, 2021; Sohan, Sai Ram ve Rami Reddy, 2024). YOLOv8, yalnızca nesne tespiti yapmakla kalmayıp; aynı zamanda görüntü sınıflandırma, insan vücut duruşunu tahmin etme, tekil nesnelere ayırt etme ve yönü belli olan nesnelere tanıma gibi birçok farklı alanda da kullanılabilir. Bunun yanı sıra, görüntüdeki anlamlı bölgeleri tespit etmeye yönelik özel bir model de bu sürümle birlikte tanıtılmıştır. Bu model, yüksek başarı elde etmesinin yanı sıra hızlı ve verimli çalışmasıyla da öne çıkmaktadır.

Modelin yapısı, önceki sürüm olan YOLOv5 ile benzer bir temel üzerine kurulmuştur; ancak daha iyi sonuçlar elde edebilmek için bazı iç düzenlemeler yapılmıştır. Bu sayede model, görüntülerdeki önemli ayrıntıları daha iyi yakalayabilmekte ve doğruluğunu artırmaktadır. Ayrıca, veriyi işlerken daha sade bir yapı kullanması sayesinde hem gereken işlem gücü azalmakta hem de modelin hızı artmaktadır.

YOLOv8'in başarısı, sadece yapısal yeniliklerle sınırlı değildir. Model, COCO ve Roboflow 100 gibi yaygın olarak kullanılan veri kümeleri üzerinde test edilmiş ve oldukça yüksek doğruluk oranlarına ulaşmıştır (Sarda, Dixit ve Bhan, 2021). Örneğin, COCO veri kümesindeki bir test setinde, YOLOv8'in büyük boyutlu modeli, bir önceki sürüm olan YOLOv5'e göre daha yüksek başarı sağlamıştır. Ayrıca, güçlü donanımlarla test edildiğinde saniyede yüzlerce görüntüyü işleyebilecek kadar hızlı olduğu görülmüştür.

Bu yeni sürümde dikkat çeken bir diğer önemli değişiklik ise, nesnelere belirleme yöntemidir. Önceki modellerde nesnelere yer, önceden tanımlanmış kutulara göre tahmin ediliyordu. Ancak YOLOv8, bu yöntemi bırakıp doğrudan nesnelere konumunu tahmin etmeye yönelmiştir. Bu yaklaşım, modelin daha az işlem yaparak daha hızlı ve doğru sonuçlar üretmesine olanak tanımıştır (Shoman vd., 2024). Ayrıca, modelin eğitimi sırasında uygulanan veri artırma yöntemleri sayesinde, farklı

ortamlarda ve görüntü düzenlerinde bile nesnelere tanıma yeteneği önemli ölçüde artırılmıştır.

Bu genel amaçlı nesne tespit modeline ek olarak, YOLOv8'in özel olarak segmentasyon görevleri için geliştirilmiş bir versiyonu da bulunmaktadır. YOLOv8-seg, YOLO ailesinin en optimize edilmiş ve hafif modellerinden biridir; küçük nesnelere tespitinde yüksek doğruluk sunmaktadır. Basitleştirilmiş mimarisi sayesinde gerçek zamanlı segmentasyon yapabilir ve düşük gecikme ile düşük hesaplama yükü gerektirir. Bu özellikleri sayesinde, hedef yapının küçük, dinamik ve hareket kaynaklı bozulmalara duyarlı olduğu göz segmentasyonu için umut vaat eden bir adaydır (Terven, Córdoba-Esparza ve Romero-González, 2023).

4.1.2. YOLOv11 Segmentasyon Modeli

YOLO11 (Jocher & Qiu, 2024), Ultralytics tarafından geliştirilen ve YOLO serisinin en yeni sürümlerinden biridir. Özellikle YOLOv8 üzerine yapılan iyileştirmelerle geliştirilmiş olan bu sürüm, farklı kullanım senaryolarına uygun beş farklı model seçeneğiyle (en küçükten en büyüğe) sunulmaktadır. Tıpkı önceki sürüm gibi, YOLO11 de nesne tespiti, görüntü sınıflandırma, insan vücut duruşu tahmini, nesne bölütleme ve yönü belli olan nesnelere tanınması gibi birçok farklı alanda kullanılabilir.

YOLO11'in öne çıkan yeniliklerinden biri, modelin görüntülerdeki bağlamsal bilgileri daha iyi anlamasını sağlayan yeni bir yapı kullanmasıdır. Bu sayede model, özellikle küçük ya da birbirine yakın nesnelere daha doğru şekilde ayırt edebilmektedir.

Ayrıca YOLOv8'de yer alan bir yapı, bu sürümde daha verimli bir yapıyla değiştirilmiştir. Yeni yapı, işlemleri daha küçük ve etkili adımlarla yaparak hem doğruluğu korumakta hem de modelin daha hızlı çalışmasını sağlamaktadır (Jegham vd., 2024).

YOLO11'in yalnızca nesne tespitiyle sınırlı kalmayıp, segmentasyon görevlerine uyarlanmış bir versiyonu da bulunmaktadır. YOLOv11-seg, geliştirilmiş özellik çıkarımı sağlayan daha derin bir ağ mimarisi sunar ve bu sayede kenar tespiti ve segmentasyon doğruluğu artar. Düşük kontrastlı görüntülerde daha istikrarlı performans sergilemekle birlikte, YOLOv8'e kıyasla daha fazla hesaplama kaynağı ve eğitim süresi gerektirir. Buna rağmen, üstün sınır tespiti yetenekleri, pupil veya iris

gibi küçük ve detaylı yapıların segmentasyonu için uygun bir seçenek haline getirmektedir (Khanam ve Hussain, 2024; Sapkota ve Karkee, 2024).

4.1.3. YOLOv12 Segmentasyon Modeli

YOLOv12, YOLO serisinin en yeni ve gelişmiş sürümüdür (Tian vd., 2025). Özellikle hız ve doğruluğu aynı anda artırmak amacıyla geliştirilen bu model, dikkat odaklı bir tasarım anlayışıyla öne çıkmaktadır. Tıpkı önceki sürümlerde olduğu gibi, en küçükten en büyüğe doğru beş farklı model seçeneği sunulmakta ve bu sayede farklı kullanım alanlarına kolayca uyum sağlayabilmektedir. Nesne tespiti, nesne bölütleme ve yönü belirli nesnelere tanıma gibi birçok uygulamada kullanılabilir.

YOLOv12'de yeni bir dikkat mekanizması kullanılmaktadır. Bu sayede model, daha az işlem gücüyle çalışmasına rağmen geniş bir görüş alanını koruyarak yüksek doğruluk sağlamaktadır. Ayrıca, modelin eğitimi sırasında daha kararlı bir öğrenme süreci geçirmesi ve daha hızlı sonuca ulaşması için geliştirilen yeni bir yapı da bu sürümde yer almaktadır.

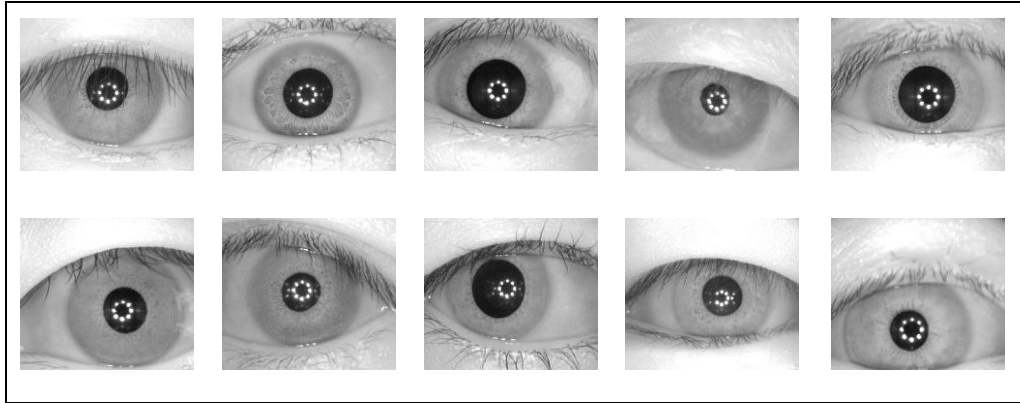
YOLOv12'nin hız konusundaki en önemli avantajlarından biri de, belleğe erişim sırasında oluşan zaman kaybını azaltan özel bir teknik kullanmasıdır. Bu sayede, geleneksel yöntemlerle çalışan modellere göre çok daha hızlı çalışır. Modelin yapısında bazı sadeleştirmeler yapılarak, karmaşıklık azaltılmıştır. Örneğin, önceki sürümlerde üst üste eklenen bazı katmanlar, bu sürümde tek bir yapı ile değiştirilmiştir. Ayrıca, daha verimli çalışması için bazı temel bileşenler sadeleştirilmiş ve hız-doğruluk dengesi açısından son derece başarılı bir sonuç elde edilmiştir.

Bu genel yapının yanı sıra, YOLOv12'nin segmentasyon görevlerine özel olarak geliştirilmiş bir versiyonu da bulunmaktadır. YOLOv12-seg, YOLO ailesinin en yeni gelişmelerinden biridir ve daha derin evrişim katmanları ile gelişmiş dikkat mekanizmalarını entegre etmektedir. Özellikle küçük nesne tespiti ve segmentasyonu görevlerinde doğruluk ve işlem hızı açısından önemli iyileştirmeler sunar. YOLOv12-seg, hem eğitim hem de çıkarım (inference) aşamalarında yüksek performansı korurken hesaplama gereksinimlerini azaltacak şekilde optimize edilmiştir (Tian, Ye ve Doermann, 2025).

4.1.4. Ground Truth Çıkarma

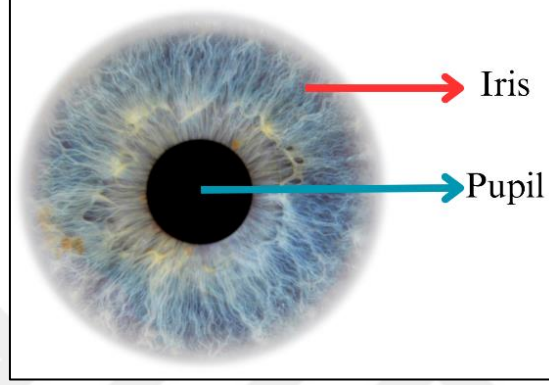
Segmentasyon temelli derin öğrenme modellerinin güvenilir bir şekilde eğitilebilmesi ve değerlendirilebilmesi, yalnızca modelin yapısal özelliklerine değil, aynı zamanda kullanılan referans verilerin doğruluğuna da bağlıdır. Bu bağlamda, modelin tahmin ettiği çıktılarla karşılaştırmak üzere kullanılan, gerçeği en iyi şekilde temsil ettiği kabul edilen etiketli veriler ground truth olarak adlandırılmaktadır. Ground truth veriler, modelin öğrenme sürecinde hedefin ne olduğunu açıkça tanımlarken; test aşamasında ise modelin başarımını ölçmek için karşılaştırma ölçütü olarak görev yapar. Özellikle segmentasyon gibi piksel düzeyinde doğruluk gerektiren görevlerde ground truth'un kalitesi, doğrudan model performansını etkileyen temel bir unsurdur. Bu nedenle, çalışmalarda kullanılan ground truth verilerin nasıl üretildiği, hangi yöntemlerle oluşturulduğu ve ne ölçüde güvenilir olduğu, model sonuçlarının yorumlanmasında büyük önem taşımaktadır.

Bu çalışmada kullanılacak verilerin bir kısmı, Çin Bilimler Akademisi Otomasyon Enstitüsü (CASIA) tarafından toplanan CASIA-Iris-Interval veri kümesine dayanmaktadır (Institute of Automation, Chinese Academy of Sciences [CASIA], 2025). Bu doğrultuda ground truth verileri, CASIA-Iris-Interval veri kümesindeki görüntülerden elde edilmiştir. YOLO segmentasyon modellerinin eğitimi için kullanılan örnek iris görüntüleri Şekil 4.1'de gösterilmiştir.



Şekil 4.1: CASIA-Iris-Interval veri kümesine ait örnek göz bebeği görüntüleri.

CASIA-Iris-Interval veri kümesi, 249 farklı kişiden toplanmış toplam 2.655 göz görüntüsü içermektedir. Her bir görüntü 320×280 piksel çözünürlüğe sahiptir ve yüksek kaliteli iris segmentasyonunu desteklemek amacıyla özel olarak derlenmiştir. Veri kümesi, kapalı ortamda ve yakın kızılötesi (NIR) aydınlatma altında, yüksek çözünürlüklü bir iris kamerası kullanılarak elde edilmiştir. Dairesel bir NIR LED dizisi, çevresel ışıklandırma ve diğer dış etkenlerin etkisini en aza indirirken iris dokusunun net şekilde görselleştirilmesini sağlamaktadır.



Şekil 4.2: Gözün anatomik yapısını gösteren ve segmentasyon işleminde hedeflenen iris (dış çember) ile pupil (iç çember) bölgeleri.

YOLO modelleriyle göz segmentasyonu gerçekleştirebilmek için, eğitim veri kümesinin doğru şekilde segmentlenmiş pupil veya iris bölgelerini içermesi gereklidir. Gözün anatomik yapısı hem irisi hem de göz bebeğini kapsadığından, bu çalışmada her iki bileşenin eşzamanlı olarak segmentlenmesi hedeflenmiştir. Segmentasyon sırasında odaklanılan özgül göz bölgelerini gösteren görsel bir temsil Şekil 4.2’de sunulmuştur.

CASIA-Iris-Interval veri kümesinde manuel etiketleme ihtiyacını ortadan kaldırmak ve yüksek kaliteli segmentasyon sağlamak amacıyla, literatürde doğruluğu kanıtlanmış bir segmentasyon yöntemi benimsenmiştir. Özellikle, A Ground Truth for Iris Segmentation başlıklı ve Halmstad Üniversitesi tarafından geliştirilmiş olan, Uluslararası Desen Tanıma Konferansı (ICPR)’da sunulan segmentasyon algoritması kullanılmıştır (Alonso-Fernandez ve Bigun, 2015; Hofbauer vd., 2014; Halmstad University Biometrics, 2025). Bu yöntemin herkese açık uygulaması kullanılarak, CASIA-Iris-Interval veri kümesindeki her bir görüntü için otomatik olarak doğruluk temelli maske (ground truth mask) üretimi gerçekleştirilmiştir. Elde edilen bu

maskeler, YOLO modellerinin eğitimi için etiketli bir veri kümesi hâline getirilmiş ve göz bebeği segmentasyon performansını artırmak amacıyla optimize edilmiştir.

Seçilen YOLO model varyantları, ilk olarak maskeleri çıkarılmış CASIA-Iris-Interval veri kümesi üzerinde eğitilmiş ve ardından zorlu koşullar altındaki BPPV hastalarına ait video kayıtları üzerinde test edilmiştir. Modellerin performansı, bir dizi değerlendirme ölçütü kullanılarak analiz edilmiştir. Segmentasyon doğruluğunu nicel olarak değerlendirmek amacıyla mean Average Precision (mAP) metriği kullanılmıştır. mAP, nesne tespiti ve segmentasyon görevlerinde en çok kullanılan değerlendirme ölçütlerinden biridir. Bu metrik, modelin farklı eşik değerlerinde yaptığı doğru ve yanlış tahminleri dikkate alarak genel başarıyı ölçer. Genellikle, her sınıf için hesaplanan ortalama kesinlik değerlerinin (Average Precision) ortalaması alınarak elde edilir (Everingham vd., 2010). Bu çalışmada, COCO değerlendirme protokolüne uygun olarak mAP@0.5:0.95 metriği kullanılmıştır. Buna ek olarak, modellerin hızı, parametre sayısı ve hesaplama maliyeti gibi unsurlar da dikkate alınarak kapsamlı bir performans analizi gerçekleştirilmiştir.

4.1.5. YOLO Model Eğitimi

Üretilen doğruluk temelli maskeler, YOLO modellerinin eğitimi için kullanılmış ve eğitim süreci, doğru göz segmentasyonu sağlamak üzere optimize edilmiştir. Gerçek zamanlı uygulamalarda hız ve doğruluk arasında denge kurabilme yetenekleri nedeniyle YOLOv8, YOLOv11 ve YOLOv12 mimarileri tercih edilmiştir. Eğitim sürecinde, CASIA-Iris-Interval veri kümesinden elde edilen göz segmentasyon maskeleri YOLO etiketleme formatına dönüştürülmüş ve giriş verisi olarak yapılandırılmıştır.

Üç YOLO modeli; çıkarım (inference) hızı, parametre sayısı, hesaplama karmaşıklığı (FLOPs) ve segmentasyon doğruluğu açısından karşılaştırılmıştır. Adil bir karşılaştırma sağlamak amacıyla, tüm modeller aynı hiperparametreler ile, sabit sayıda epoch ve sabit giriş çözünürlüğü kullanılarak eğitilmiştir. Eğitim süreci boyunca tutarlı bir veri kümesi kullanılması, nesnel bir performans değerlendirmesine olanak tanımıştır.

Eğitimin tamamlanmasının ardından YOLO modelleri, göz bebeği ve iris tespitinde başarılı sonuçlar vermiş ve BPPV tanısıyla ilişkili zorlu senaryolarda test edilmiştir. Bu değerlendirme aşamasında, eğitilmiş modeller BPPV video veri kümesine uygulanmış; segmentasyon çıktıları görsel inceleme ve nitel karşılaştırma yoluyla analiz edilmiştir. Ayrıca, YOLO modelleri tarafından üretilen sonuçlar, SAM2 modeli ile karşılaştırılarak her iki yöntemin farklı koşullar altındaki performansı değerlendirilmiştir.

4.2. Segment Anything Model 2

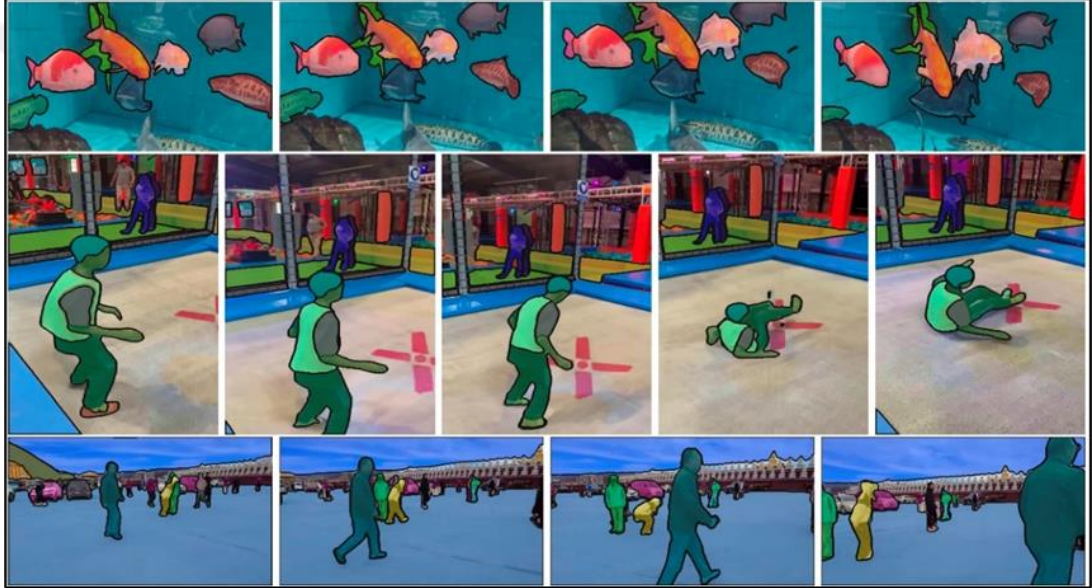
Segment Anything Model (SAM), bilgisayarla görme ve derin öğrenme alanında segmentasyon zorluklarını aşmak amacıyla Meta AI tarafından genel amaçlı bir çözüm olarak geliştirilmiştir (Meta AI Research, 2025). Nesne sınıfından bağımsız olarak segmentasyon yapabilecek şekilde tasarlanan SAM, büyük ölçekli veri kümeleri üzerinde eğitilmiş olup, farklı görüntü türleri arasında güçlü bir genelleme yeteneğine sahiptir (Kirillov vd., 2023). Bu başarımının temelinde, kullanıcıdan aldığı girişlere dayalı esnek ve etkileşimli çalışma prensibi yer almaktadır.

SAM, kullanıcı tarafından sağlanan girişlere (örneğin tıklama, sınırlayıcı kutu veya metin) dayalı olarak herhangi bir nesneyi görüntü içinde segmentleyebilme yeteneğine sahiptir. Bu yönüyle, sıfır örneklili segmentasyon (zero-shot segmentation) kavramını pratikte uygulanabilir hâle getirmiştir (Kirillov vd., 2023). Ancak SAM'in özellikle video verilerinde zaman boyutunu dikkate almaması ve yüksek çözünürlükte tutarsız sonuçlar üretmesi, daha gelişmiş bir modele ihtiyaç doğurmuştur. Bu ihtiyaçların bir sonucu olarak geliştirilen SAM2, hem mimari düzeyde hem de işlevsel olarak önemli yenilikler içermektedir. SAM2'nin temel farkı, görüntü segmentasyonu ile sınırlı kalmayıp video segmentasyonu gibi zaman boyutlu verilere de etkin şekilde uyulanabilmesidir.

SAM2, önceki sürümdeki maske çıkarıcı (mask decoder) yapısını geliştirerek, daha kararlı ve hassas segmentasyon sonuçları üretebilmekte; aynı zamanda düşük hesaplama yüküyle daha yüksek doğruluk sunmaktadır (Zhang ve Tang, 2025). Görüntü segmentasyonunda SAM2, farklı çözünürlüklerde güçlü genel performans gösterirken; video segmentasyonu görevlerinde ise zaman içindeki nesne tutarlılığını koruyarak hareketli nesnelerin daha sağlıklı izlenmesini mümkün kılar. Özellikle SAM2'nin video segmentasyonundaki başarısı, önceki modellerin çerçeve bazlı

çalışmasından kaynaklanan tutarsızlık sorunlarını büyük ölçüde azaltmasıyla dikkat çekmektedir.

Yapay zekâ tabanlı modellerin genellikle tek bir görüntü üzerinden çalışması, zaman boyutu içeren görevlerde önemli performans kayıplarına yol açarken; SAM2 bu sorunu yeni dikkat mekanizmaları ve daha esnek encoder yapısıyla minimize etmektedir. Ayrıca SAM2, geniş çapta önceden eğitilmiş veri temeliyle sıfır örnekleli kullanım senaryolarında da güçlü sonuçlar üretmektedir. Bu gelişmeler ışığında, SAM2 yalnızca görüntüde değil, özellikle tıbbi video analizleri, hareketli nesne takibi ve semantik video segmentasyonu gibi alanlarda da umut vaat eden bir çözüm olarak öne çıkmaktadır (Zhang ve Tang, 2025).



Şekil 4.3: SAM2 uygulanan veri kümesinden örnek video kareleri.

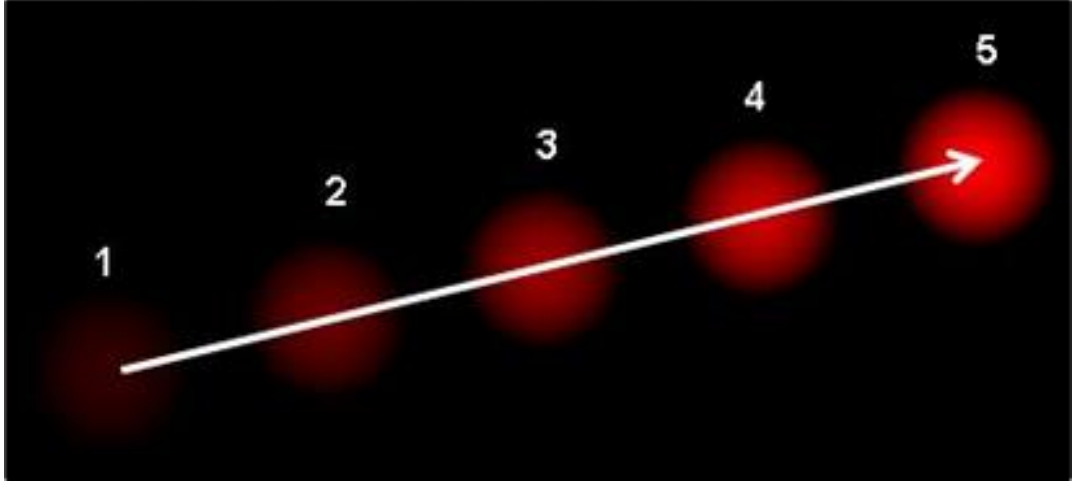
Kaynak: Kirillov vd., 2023.

Tüm bu gelişmeler üzerine mimarisi daha da optimize edilen SAM2, segmentasyon performansını artırmak ve işlem süresini azaltmak amacıyla ileri düzeyde yapılandırılmıştır. Bu temelin üzerine inşa edilen SAM2, segmentasyon performansını artırmak ve işlem süresini azaltmak amacıyla geliştirilmiştir. Gerçek zamanlı uygulamalar için özel olarak optimize edilen SAM2, hem tıbbi hem de endüstriyel alanlarda yüksek hassasiyet gerektiren segmentasyon görevleri için etkili bir çözüm sunmaktadır (Ravi vd., 2024; Chen vd., 2024; Xiong vd., 2024).

SAM2'nin önceki sürümüne kıyasla en önemli iyileştirmelerinden biri, yaklaşık altı kat daha hızlı çalışabilmesidir. Bu hız artışı, optimize edilmiş bellek kullanımı ve paralel işlem mimarisi sayesinde elde edilmiştir. Geliştirilmiş özellik çıkarım mekanizmaları, rafine edilmiş ağ mimarisi ve küçük nesnelere segmentasyonundaki başarımı, yüksek doğruluğun korunurken düşük hesaplama maliyetiyle çalışmasını mümkün kılmaktadır (Videnovic vd., 2025; Zhou vd., 2025). Bu gelişme, özellikle karmaşık ve küçük anatomik yapıların segmentasyonunun kritik olduğu tıbbi görüntüleme alanında son derece değerlidir. Nitekim, SAM2 modeli bu alanda yüksek güvenilirlik göstermiştir (Ravi vd., 2024; Ma vd., 2024, Li vd., 2024).

4.3. Optik Akış Yöntemi

Video işlemede en önemli konulardan biri, çerçeveler (frame) arasındaki değişimleri doğru bir şekilde yakalayabilmektir. Seçilen bir çerçevedeki bir noktanın bir sonraki çerçevede de olup olmadığı ya da nerede olduğu bilgisi, hareketin tespiti ve analizi açısından kritiktir. Bu tür analizleri gerçekleştirebilmek için optik akış (optical flow) yöntemi yaygın olarak kullanılmaktadır.



Şekil 4.4: Ardışık hareket eden bir topun yer değiştirme vektörünü gösteren ok.

Kaynak: OpenCV Documentation, Optical Flow. Erişim: 20 Nisan 2025.

Optik akış, bir görüntü dizisi boyunca zaman içinde meydana gelen hareketin yönünü ve büyüklüğünü tahmin etmek için kullanılan bir görüntü işleme tekniğidir (Horn ve Schunck, 1981). İnsan görsel sisteminin, bir nesnenin hareketini tanımak için genellikle o nesnenin konumundaki değişiklikleri kullanması gibi, optik akış da ardışık görüntülerdeki piksel hareketlerini analiz eder. Bu teknik sayesinde hem kamera hareketleri hem de sahnedeki nesnelere hareketleri takip edilebilir. Optik akış, bu

kavrama dayanarak, bir kameranın bakış açısındaki bir nesnenin zaman içindeki konumundaki deęişiklikleri ölçer. Optik akış, birçok uygulama alanında kullanılır:

- **Hareket Algılama ve İzleme:** Video izleme sistemlerinde, nesnelerin hareketini izlemek ve tespit etmek için kullanılır (Chen vd., 2023).
- **Hareket Tabanlı Navigasyon:** Robotların veya otonom araçların çevresindeki nesnelere algılamak ve takip etmek için kullanılabilir (Rashed vd., 2019).
- **Eylem Tanıma:** Bir video içindeki insan eylemlerini tanımak ve sınıflandırmak için kullanılabilir (Afza vd., 2021).

Optik akışın ana prensipleri aşağıdaki gibidir.

- **Hareket Algılaması:** Optik akış, bir video içindeki piksellerin veya görüntü noktalarının zaman içindeki hareketini tanımlar. Bu hareket, kameranın hareketinden kaynaklanıyor olabilir veya sahnedeki nesnelerin hareketinden kaynaklanıyor olabilir.
- **Piksel Deęişimleri:** Optik akış, her bir pikselin bir karedeki konumunu diğer karelere göre izleyerek hareket eden nesnenin veya kameranın hareketine dair bir tahmin oluşturur. Bu, her pikselin zaman içindeki konumundaki deęişikliklerin vektörel bir gösterimidir.
- **Optik Akış Vektörleri:** Optik akış, her bir pikselin hareketini tanımlamak için optik akış vektörleri adı verilen vektörler kullanır. Bu vektörler, bir karedeki bir pikselin bir sonraki karedeki konumuna olan deęişimini belirtir.
- **Hesaplama Yöntemleri:** Optik akış vektörleri, çeşitli hesaplama yöntemleri kullanılarak elde edilir. Bu yöntemler arasında Lukas-Kanade yöntemi, yoğun optik akış yöntemi, Horn-Schunck yöntemi, veya daha modern derin öğrenme tabanlı yöntemler bulunabilir.

Optik akış algoritmaları genellikle hareket kestirimi, piksel yoğunluğu deęişimleri ve vektör alanı hesaplaması ilkelerine dayanır. Her bir pikselin zaman içindeki konum deęişimi, matematiksel olarak optik akış vektörleri ile temsil edilir. Bu vektörler, hareketin yönünü ve hızını tanımlar. Bu tür temsiller, dinamik öğelerin takibini ve zamansal deęişimlerin modellenmesini mümkün kılar (Afza vd., 2021; Chen vd., 2023; Rashed vd., 2019). Optik akış yöntemi, başlıca iki ana yaklaşımla uygulanabilir: seyrek (sparse) ve yoğun (dense).

Seyrek optik akış, yalnızca belirli anahtar noktalara (örneğin köşe veya kenar gibi özellik noktalarına) odaklanır ve bu noktaları kareler boyunca takip eder. Bu yöntemin bilinen bir örneği Lucas–Kanade yöntemidir. Bu yöntem, bir görüntüdeki parlaklık değişikliklerinin o görüntüdeki nesnelerin hareketine nasıl bağlı olduğunu modellemektedir.

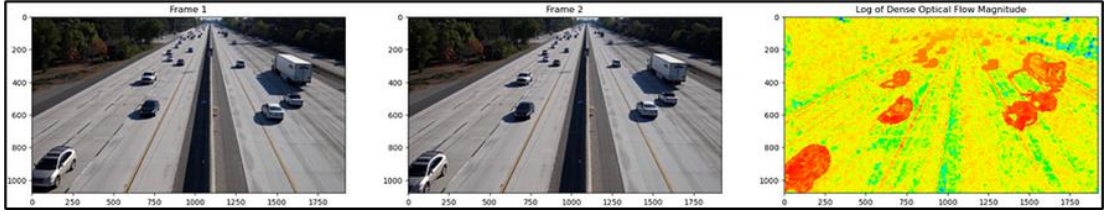
Lucas-Kanade yöntemi, bir nokta etrafında 3x3'lük bir bölge (kernel) kullanır. Yani belirlenen bir nokta ve çevresindeki 9 komşu pikselin aynı harekete sahip olduğu varsayılır. Küçük ölçekli hareketler bu kernel ile çözülebilirken, büyük hareketlerde piramit (pyramid) yapısı kullanılarak çözüm sağlanır. Piramidin üst seviyelerinde küçük hareketler baskılanır ve büyük hareketler küçük hareketlere dönüştürülerek daha stabil hesaplamalar yapılır (Lucas ve Kanade, 1981).



Şekil 4.5: Lukas-Kanade yöntemi uygulanmış görseller.

Kaynak: OpenCV Documentation, Optical Flow. Erişim: 20 Nisan 2025.

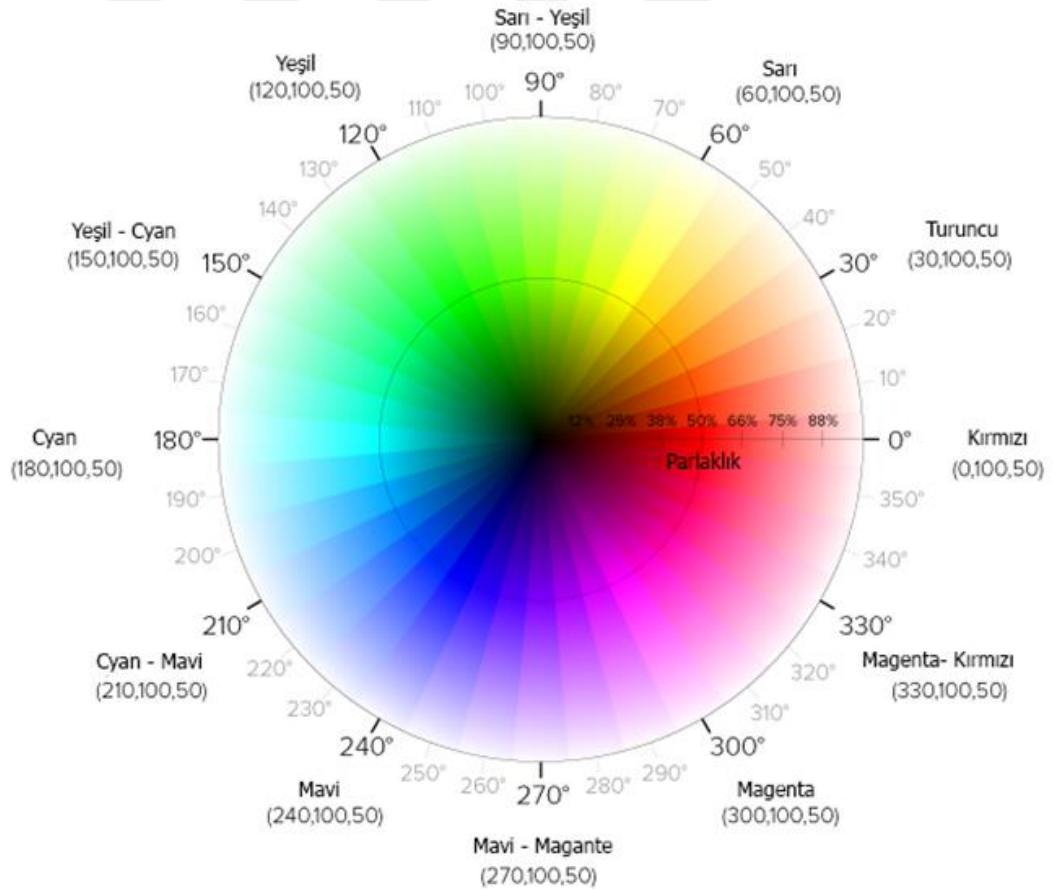
Lucas-Kanade yöntemi küçük yer değiştirmelerde başarılı olsa da, hızlı veya büyük ölçekli hareketlerde yetersiz kalabilir. Yoğun (dense) optik akış, görüntüdeki her piksel için hareket vektörleri hesaplayarak daha ayrıntılı bir analiz sunar. Bu yaklaşımın önemli bir temsilcisi olan Farneback yöntemi, tüm çerçeve boyunca piksel bazlı hareket alanları üretir ve bu sayede hızlı hareketler ile çevresel gürültülere karşı daha dayanıklı hale gelir (Farneback, 2003).



Şekil 4.6: Sıralı video kareleri ve yoğun optik akışları.

Kaynak: OpenCV Documentation, Optical Flow. Erişim: 20 Nisan 2025.

Farneback algoritması, ardışık görüntü karelerini küçük bloklara (örneğin 4x4 piksel) ayırarak her blok için bir hareket vektörü üretir. Bu işlem, çerçevedeki tüm noktalar için gerçekleştirilerek yoğun bir optik akış haritası oluşturur. Kullanım alanları arasında hareket algılama, nesne takibi ve video sabitleme gibi pek çok uygulama bulunmaktadır. Bu çalışmada, göz bebeği segmentasyon sonuçları üzerinde nistagmusun yönünü ve şiddetini belirlemek amacıyla Farneback Yoğun Optik Akış algoritması kullanılmıştır.



Şekil 4.7: HSV uzayındaki renk kodları

Kaynak: Pişkin, M., Erişim: 20 Nisan 2025.

Elde edilen hareket vektörleri, görsel olarak daha sezgisel bir biçimde yorumlanabilmeleri için HSV (Hue, Saturation, Value) renk uzayına dönüştürülerek görselleştirilmiştir (Gonzales ve Wintz, 1987). Bu renk modeli algısal olarak sezgiseldir ve hareket dinamiklerinin net biçimde yorumlanmasını sağlar; burada ton (hue) hareketin yönünü, parlaklık (value) ise hareketin şiddetini temsil eder (Smith, 1978).

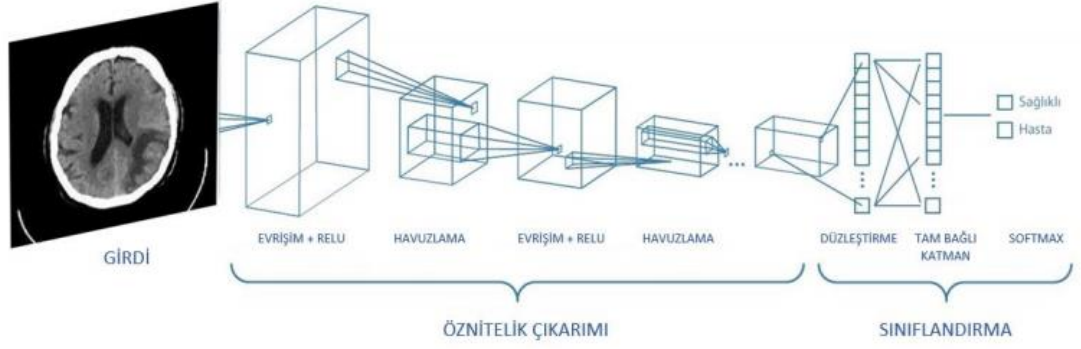
Optik akış vektörleri, ilk olarak yoğun Farneback yöntemi ile hesaplanmış ve ardından HSV renk uzayına dönüştürülmüştür. Bu dönüşüm, nistagmusla ilişkili hareket örüntülerinin yön ve şiddetini daha iyi anlamak ve sezgisel şekilde görselleştirmek amacıyla gerçekleştirilmiştir.

4.4. Sınıflandırma

Göz hareketi yönünü sınıflandırmak amacıyla, optik akış verilerini hem mekânsal hem de zamansal boyutlarda ortaklaşa analiz eden bir derin öğrenme mimarisi tasarlanmıştır. Bu mimaride, Convolutional Neural Network (CNN) mekânsal özellikleri çıkarmak için, Gated Recurrent Unit (GRU) ise ardışık kareler arasındaki zamansal ilişkileri modellemek için kullanılmıştır. Model, iris bölgesine ait optik akış tensörlerinden mekânsal özellikleri çıkararak ve ardışık kareler boyunca zamansal bağımlılıkları öğrenerek hareket örüntülerini öğrenmeyi hedeflemektedir.

4.4.1. Convolutional Neural Networks

CNN, LeCun vd. (1989) tarafından 1989 yılında önerilen ileri beslemeli bir yapay sinir ağı mimarisidir. CNN'ler, giriş verilerinden karmaşık ve hiyerarşik özellikleri otomatik olarak öğrenebilen derin öğrenme mimarileridir (Zareapoor, Shamsolmoali ve Yang, 2019). Ardışık konvolüsyon, havuzlama (pooling) ve aktivasyon işlemleri sayesinde; görüntü sınıflandırma, nesne tanıma, segmentasyon ve tespit gibi birçok görevde yaygın olarak kullanılmaktadır (Shamsolmoali, Li ve Wang, 2019; Basha, Dubey, Pulabaigari ve Mukherjee, 2020).



Şekil 4.8: Örnek CNN yapısı.

Kaynak: Özcan, K., Medium, Erişim: 20 Mayıs 2025.

Sinir ağları genel anlamda makinelerin, insanlar gibi öğrenmesini hedefler. CNN'ler, bu bağlamda öğrenme sürecinde özellikle görsel veriler üzerinde oldukça etkili sonuçlar sunan özel bir sinir ağı türü olarak öne çıkar (LeCun vd., 1998). Bu ağlar, gözetimli öğrenme yöntemlerini temel alır; yani ağ, eğitim sırasında etiketlenmiş verilerden öğrenerek belirli sınıflar arasında ayrım yapmayı öğrenir. Örneğin, eğitim verisinde yalnızca "kedi", "köpek" ve "kuş" gibi genel sınıflar bulunuyorsa, ağ bu sınıflar arasında ayrım yapmayı öğrenir. Ancak, eğer daha ayrıntılı türlere ait örnekler sağlanırsa, ağ bu ince ayrımları da öğrenebilir.

CNN mimarisi, genellikle bir giriş katmanı, bir çıkış katmanı ve arada yer alan birden çok gizli katmandan oluşur. Gizli katmanların sayısı ve türleri, ağın çözmesi beklenen probleme bağlı olarak değişkenlik gösterebilir. Bu mimaride temel olarak, giriş görüntüsünden özellik çıkarımı yapılır ve bu işlem yerel filtrelerin görüntü üzerinde kaydırılmasıyla gerçekleştirilir. Bu süreç sonunda elde edilen özellik haritaları, boyutları küçültülerek sonraki katmanlara iletilir. Bu işlemler zinciri, anlamlı özellikler çıkarılincaya dek devam eder ve en sonunda bu özelliklere dayalı bir sınıflandırma işlemi gerçekleştirilir (Ravi vd., 2017).

CNN mimarileri genellikle üç temel yapı taşından oluşur: konvolüsyon katmanları, havuzlama katmanları ve tam bağlantılı katmanlar (Schmidhuber, 2015). Özellik çıkarım süreci, konvolüsyon ve havuzlama katmanlarıyla gerçekleştirilirken, tam bağlantılı katmanlar genellikle sınıflandırıcı olarak görev yapar (Labib vd., 2022). Bu yapının birçok başarılı örneği bulunmaktadır; AlexNet (Krizhevsky vd., 2012), OverFeat (Sermanet vd., 2013) ve GoogLeNet (Szegedy vd., 2015) bu yapıya dayalı olarak geliştirilmiş önemli CNN mimarileri arasında yer alır.

Konvolüsyon katmanları, giriş verisine belirli işlevleri uygulayan filtreler (çekirdekler) içerir. Bu filtreler genellikle sabit boyutludur ve öğrenme sürecinde optimize edilirler (Ravi vd., 2017). Her filtre, giriş görüntüsünün üzerinde belirli adımlarla (stride) kaydırılarak uygulanır ve her konumda bir aktivasyon değeri üretir. Bu şekilde, her bir nöron yalnızca önceki katmandaki sınırlı sayıda nöronla bağlantılıdır; bu sınırlı alan, "alıcı alan" (receptive field) olarak tanımlanır. Her filtre, tüm görüntü boyunca aynı ağırlık ve sapma değerlerini kullanır; bu mekanizma "ağırlık paylaşımı" olarak adlandırılır ve modelin parametre sayısını azaltarak verimliliği artırır (Chandrakumar ve Kathirvel, 2016; Sankar vd., 2016; Ravi vd., 2017).

Filtre sayısı ve boyutları, modelin başarısını etkileyen önemli hiperparametrelerdir. Az sayıda filtre kullanıldığında model yeterince desen tanıyamazken, fazla sayıda filtre kullanıldığında parametre sayısı ve eğitim süresi artar. Bu nedenle denge gözetilmelidir. İlk katmanlarda genellikle daha büyük filtreler tercih edilirken, sonraki katmanlarda daha küçük filtreler kullanılır. Ayrıca, giriş görüntüsünün kenarlarına sıfır değerli pikseller eklenerek yapılan "padding" işlemi, kenar bilgilerini koruyarak çıktı boyutunu sabit tutmaya yardımcı olur. Konvolüsyon işlemi sonrası her bir nöronun çıktısı aşağıdaki formülle hesaplanır (Nielsen, 2015):

$$\sigma_{0,0} = f\left(b + \sum_{t=0}^c \sum_{r=0}^c w_{t,r} i_{0+t,0+r}\right) \quad (4.1)$$

Burada f , genellikle ReLU, sigmoid veya tanh gibi doğrusal olmayan bir aktivasyon fonksiyonudur.

Konvolüsyon katmanından elde edilen çıktılar, genellikle doğrusal olmayan bir aktivasyon fonksiyonundan geçirilir. Aktivasyon fonksiyonları, sinir ağlarının doğrusal olmayan ilişkileri öğrenmesini sağlar. En sık kullanılan aktivasyon fonksiyonları arasında ReLU, sigmoid, hiperbolik tanjant ve SoftMax yer alır.

$$\text{çıkıtı} = \frac{\text{girdi} - \text{filtre} + 2 \cdot \text{padding}}{\text{stride}} + 1 \quad (4.2)$$

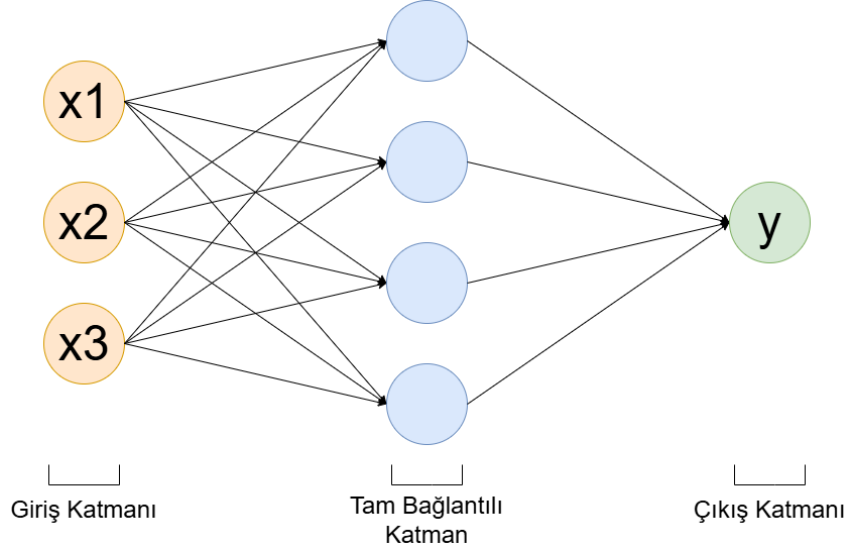
"Stride", filtrelerin giriş görüntüsü üzerinde kaç birimlik adımlarla kaydırıldığını belirler. Örneğin, stride değeri 1 olduğunda filtre görüntü üzerinde birer birimlik adımlarla kaydırılır; stride 2 olduğunda ise bu adım iki birim olur. "Padding" işlemi ise, görüntünün kenarına sıfır (0) değerli pikseller eklenerek çıktı boyutunun korunmasını sağlar.

Aktivasyon işleminden sonra, havuzlama katmanları devreye girerek özellik haritalarının boyutunu küçültür. Bu işlem, parametre sayısını ve hesaplama yükünü azaltırken aynı zamanda yerel çevirilere karşı ağırlık dayanıklılığını artırır. En yaygın havuzlama yöntemleri maksimum havuzlama, ortalama havuzlama ve L2 havuzlamadır. Maksimum havuzlama, belirli bir alandaki en yüksek değeri alırken, ortalama havuzlama ortalamayı, L2 havuzlaması ise L2 normunu hesaplar (Nielsen, 2015).

Konvolüsyon ve havuzlama işlemleri tamamlandıktan sonra elde edilen çok boyutlu özellik haritaları, düzleştirilerek tek boyutlu bir vektöre dönüştürülür. Bu vektör, tam bağlantılı katmanlara giriş olarak verilir. Bu katmanlarda her nöron, önceki katmandaki tüm nöronlarla bağlantılıdır ve çıktı aşağıdaki formülle hesaplanır:

$$f_{c_1} = f(b + \sum_{q=1}^M w_{1,q} * o_q) \quad (4.3)$$

Burada f aktivasyon fonksiyonu, M önceki katmandaki nöron sayısı, o_q bu katmandaki q . nöronun çıktısı, $w_{1,q}$ ağırlık değeri ve f_{c_1} tam bağlantılı katmandaki ilk nöronun çıktısıdır.



Şekil 4.9: Tam Bağlantılı Katman Örneği.

Tam bağlantılı katmanların sonunda genellikle bir SoftMax aktivasyon fonksiyonu kullanılır. Bu fonksiyon, çok sınıflı sınıflandırma problemlerinde sınıflara ait olasılık değerlerini üretir. SoftMax fonksiyonu aşağıdaki gibi tanımlanır:

$$class_j = \frac{\exp(sf_j)}{\sum_q \exp(sf_q)} \quad (4.4)$$

Burada $class_j$, j . sınıfa ait olasılığı; sf_j , softmax giriş vektöründeki j . elemanı; sf_q , tüm sınıf skorlarını ifade eder (Nielsen, 2015).

Sonuç olarak, CNN mimarisi, görüntü veya diğer yapılandırılmış verilerden özellikleri etkili bir biçimde çıkarmaya ve bu özelliklere dayanarak yüksek doğrulukla sınıflandırma yapmaya olanak tanıyan güçlü bir yapay sinir ağıdır.

4.4.2. Gated Recurrent Unit (GRU)

GRU (Gated Recurrent Unit), zamansal modelleme amacıyla CNN ile entegre edilmiş olup, Uzun Kısa Süreli Bellek (LSTM) mimarisine göre daha basit ve hesaplama açısından daha verimlidir. GRU, bilgi akışını zaman içinde düzenleyen iki temel kapı mekanizması içerir: güncelleme kapısı (update gate) z_t ve sıfırlama kapısı (reset gate) r_t (Athira vd., 2018; Li vd., 2021; Pan vd., 2020). Bu kapılar, önceki zaman adımından ne kadar bilginin korunacağını veya atılacağını belirler. GRU, bu kapı mekanizmaları sayesinde ardışık verilerdeki kısa vadeli ve uzun vadeli zamansal bağımlılıkları etkili bir şekilde öğrenebilir.

$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1}) \quad (4.5)$$

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1}) \quad (4.6)$$

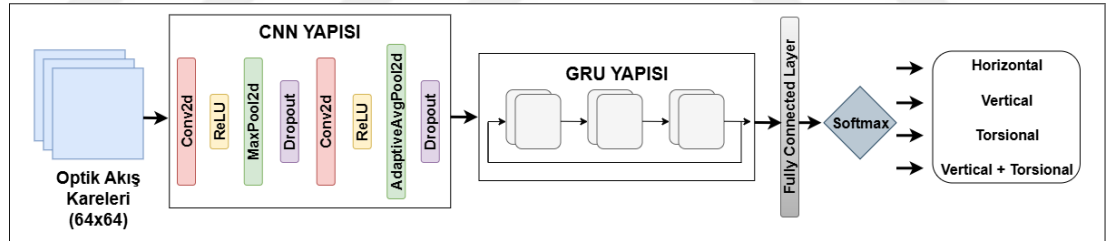
$$h_{t,i} = \tanh(r_t \circ U h_{t-1} + W x_t) \quad (4.7)$$

$$h_t = (1 - z_t) \circ h_t + z_t \circ h_{t-1} \quad (4.8)$$

Bu denklemde, $\sigma(\cdot)$ sigmoid aktivasyon fonksiyonunu, $W^{(r)}$ ve $W^{(z)}$ ise sırasıyla sıfırlama (reset) ve güncelleme (update) kapılarına ait öğrenilen ağırlık matrislerini ifade eder. \circ operatörü, bileşen bazlı (element-wise) çarpımı temsil eder; yani, iki vektörün aynı indislerdeki bileşenleri birbirleriyle çarpılır.

4.4.3. CNN-GRU Hibrit Model

Bu çalışmada, nistagmus yönü sınıflandırması için geliştirilen önerilen CNN-GRU mimarisi Şekil 4.10'da sunulmakta ve performansını en üst düzeye çıkarmak üzere yapılandırılan hiperparametreler ayrıntılı olarak açıklanmaktadır. Parametre ayarları, literatürdeki ön tanımlı değerler yerine deneysel gözlemlere dayalı olarak deneysel şekilde belirlenmiştir.



Şekil 4.10: Önerilen CNN-GRU Hibrit Model Mimarisi.

Önerilen modelde, iki kanallı optik akış görüntüleri—yatay (x) ve dikey (y) hareket bileşenlerini temsil eden—ardışık iki konvolüsyon katmanından geçirilerek mekânsal öznitelikler çıkarılmıştır. Her konvolüsyon işlemi bir ReLU aktivasyon fonksiyonu izlemiş ve boyut indirgeme işlemi maksimum havuzlama (max-pooling) ile gerçekleştirilmiştir. Elde edilen öznitelik haritaları daha sonra, sabit boyutlu bir gösterim elde etmek amacıyla uyarlanabilir ortalama havuzlama (adaptive average pooling) katmanına aktarılmıştır.

Üretilen öznitelik vektörleri, zamansal sıra hâlinde gruplanarak GRU modülüne giriş olarak verilmiştir. GRU çıktıları, zamansal temsili özetlemek amacıyla ortalanmış ve

sınıflandırma aşamasına iletilmiştir. Sınıflayıcı, ReLU aktivasyonlu iki katmanlı tam bağlantılı (fully connected) bir yapıdan oluşmakta ve aşırı öğrenmeyi (overfitting) önlemek için bir dropout katmanı içermektedir. Model, çok sınıflı çapraz entropi (multiclass cross-entropy) kayıp fonksiyonu ile eğitilmiş, optimizasyon işlemi ise Adam algoritması kullanılarak gerçekleştirilmiştir.

4.4.4. Sınıflandırma Başarı Ölçütleri

Sınıflandırma modelinin çıktılarının doğruluğunu değerlendirmek amacıyla sıklıkla kullanılan araçlardan biri karışıklık matrisidir (confusion matrix) (Sathyanarayanan ve Tantri, 2024). Bu matris, sınıflandırıcı tarafından yapılan tahminler ile gerçek sınıf etiketleri arasındaki ilişkiyi tablo halinde gösterir. Makine öğrenmesi ve istatistiksel sınıflandırma alanında, karışıklık matrisi genellikle dört temel bileşene ayrılır:

- **True Positive (TP – Doğru Pozitif):** Modelin, pozitif olan bir örneği doğru şekilde pozitif olarak sınıflandırdığı durumdur. Örneğin, hasta olan bir kişiyi “hasta” olarak tahmin etmek.
- **True Negative (TN – Doğru Negatif):** Modelin, negatif olan bir örneği doğru şekilde negatif olarak sınıflandırdığı durumdur. Örneğin, hasta olmayan bir kişiyi “sağlıklı” olarak tanımak.
- **False Positive (FP – Yanlış Pozitif):** Gerçekte negatif olan bir örneğin model tarafından pozitif olarak sınıflandırılmasıdır. Örneğin, sağlıklı bir kişiye “hasta” demek.
- **False Negative (FN – Yanlış Negatif):** Gerçekte pozitif olan bir örneğin model tarafından negatif olarak sınıflandırılmasıdır. Örneğin, hasta bir kişiyi “sağlıklı” olarak değerlendirmek.

Bu dört temel değer üzerinden çeşitli performans ölçütleri hesaplanabilir (Rainio vd., 2024). Bunlardan ilki olan doğruluk, modelin doğru sınıflandırdığı örneklerin tüm örnekler içindeki oranını verir:

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.9)$$

Ancak sınıflar arasında dengesizlik varsa, bu oran yüksek çıksa bile modelin gerçek performansı düşük olabilir. Bu nedenle kesinlik ve duyarlılık gibi ölçütler öne çıkar. Kesinlik, modelin pozitif olarak sınıflandırdığı örneklerin ne kadarının gerçekten pozitif olduğunu gösterir. Yani, “model pozitif dediğinde ne kadar güvenilir?” sorusunun yanıtıdır:

$$Kesinlik = \frac{TP}{TP + FP} \quad (4.10)$$

Duyarlılık ise, gerçek pozitif örneklerin ne kadarının doğru tespit edildiğini gösterir. Bu da “pozitif olanlar içinde model kaçını doğru buldu?” sorusuna karşılık gelir:

$$Duyarlılık = \frac{TP}{TP + FN} \quad (4.11)$$

Her iki ölçüt de farklı bakış açıları sunduğundan, bunları birleştiren F1 skoru, hem kesinlik hem de duyarlılığı dengeleyen harmonik ortalama olarak tanımlanır:

$$F1 = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (4.12)$$

F1 skoru özellikle sınıf dağılımının dengesiz olduğu durumlarda, yani yanlış negatif ve yanlış pozitiflerin etkisinin benzer olduğu senaryolarda, doğruluktan daha anlamlı bir gösterge sunar. Eğer yanlış sınıflandırma maliyetleri arasında fark varsa (örneğin bir hastayı “sağlıklı” olarak sınıflandırmak, “sağlıklı” birini “hasta” olarak sınıflandırmaktan daha riskliyse), F1 skoru yerine kesinlik veya duyarlılığın ayrı ayrı değerlendirilmesi gerekebilir.

Tıbbi görüntü analizinde, doğru segmentasyon tanı süreçleri için kritik bir öneme sahiptir. Intersection over Union (IoU), bilgisayarla görüde nesne tespiti ve segmentasyon modellerinin performansını değerlendirmek için kullanılan temel metriklerden biridir. IoU, bir modelin tahmin ettiği bölge ile gerçek (ground truth) bölge arasındaki örtüşmeyi ölçerek segmentasyon doğruluğunu hesaplar. Bu metrik, özellikle tıbbi görüntüleme segmentasyon başarımını değerlendirmek için yaygın olarak kullanılmaktadır.

IoU, tahmin edilen segmentasyon maskesi (predicted mask) ile gerçek segmentasyon maskesi (ground truth mask) arasındaki kesişim (intersection) ve birleşim (union) alanlarının oranı olarak hesaplanır:

$$IoU = \frac{TP}{(TP + FP + FN)} \quad (4.13)$$

Bu formülde:

- Kesişim (intersection), tahmin edilen segmentasyon maskesi ile gerçek bölgenin örtüştüğü alanı temsil eder.
- Birleşim (union), tahmin edilen ve gerçek segmentasyon maskelerinin toplam alanını ifade eder.

IoU, modelin segmentasyon doğruluğunu ölçerek göz bebeği segmentasyonu gibi hassas analizlerde kullanılmaktadır. Görüntü bölümlendirmesinde tahmin edilen alanlar dikdörtgen sınırlayıcı kutular (bounding boxes) yerine serbest şekilli maskeleme ile ifade edildiği için, IoU burada piksel bazlı analizle hesaplanmaktadır.



Şekil 4.11: Görüntü segmentasyonunda tahmin ve gerçek maskelerin karşılaştırılması. Model tarafından tahmin edilen segmentasyon maskesi (a), Gerçek (ground-truth) segmentasyon maskesi (b), Tahmin ve gerçek maskelerin üst üste bindirilmiş hali (c).

IoU, tahmin edilen maske veya sınırlayıcı kutu ile gerçek (ground truth) değer arasındaki örtüşmeyi nicel olarak ifade eder ve modelin tekil bir örnek üzerindeki başarımını gösterir. Ancak, yalnızca IoU değeri üzerinden yapılan değerlendirme, modelin genel doğruluk eğrisini tam olarak yansıtmayabilir. Bu nedenle, daha kapsamlı bir analiz için mAP metriği kullanılmaktadır (Lin et al., 2014).

Özellikle COCO değerlendirme kriterlerine göre hesaplanan mAP(50-95), IoU eşik değerlerini %50 ile %95 arasında 0.05 artışlarla dikkate alarak her bir eşığe karşılık gelen AP değerlerinin ortalamasını alır. Bu yöntem, yalnızca yüksek örtüşmelerde değil, farklı hassasiyet seviyelerinde de modelin tutarlılığını ölçmeye olanak tanır. Böylece, modelin hem düşük hem de yüksek IoU eşiklerinde gösterdiği genel başarı düzeyi daha bütüncül bir şekilde değerlendirilebilir.

Bu çalışmada, BPPV tanısına yönelik geliştirilen görüntü tabanlı sistem kapsamında gözbebeği segmentasyonu için YOLO ve SAM2 tabanlı modeller kullanılmıştır. YOLO tabanlı modeller için segmentasyon performansı, modelin varsayılan değerlendirme yöntemi olan mAP(50-95) metriği üzerinden nicel olarak değerlendirilmiştir. Segmentasyon yöntemlerinin başarısı, nihai sınıflandırma performansına katkıları doğrultusunda karşılaştırmalı olarak analiz edilmiştir.

4.5. Açıklanabilir Yapay Zeka

Yapay zekâ, artık hayatımızın pek çok alanında yerini almış durumda. Sağlık, finans, güvenlik, ulaşım gibi birçok sektörde karşılaşılan zorlu sorunların çözülmesine yardımcı oluyor ve geleneksel yöntemlerin yerini alıyor (Ali vd., 2023). Örneğin, akıllı telefonlarımızdaki uygulamalarda farklı görevleri yerine getiriyor (Georgiev vd., 2017), otomobillerde kazaları önlemeye katkı sağlıyor (Jain vd., 2015), bankalarda yatırım ve kredi kararlarının alınmasına yardımcı oluyor (Chong, Han ve Park, 2017; Pham ve Shen, 2017), hastanelerde doktorların hastalıkların tespit ve teşhisinde destek almasını sağlıyor (Nie vd., 2015). Polis teşkilatlarında delil toplama ve işlemleri kolaylaştırma (Goswami vd., 2014), ordularda stratejik kullanım (Lundén ve Koivunen, 2016) ve sigorta sektöründe risk analizi (Dong vd., 2016) gibi alanlarda da etkin şekilde kullanılıyor. Üstelik yüksek performansı sayesinde, birçok kurum yapay zekâyı iş süreçlerine entegre etmeye çalışıyor; çünkü artık bu sistemler, pek çok görevde insan performansına yakın ya da ondan üstün sonuçlar verebiliyor (Enholm vd., 2022).

Yapay zekânın en güçlü yönlerinden biri, veriye dayalı kararlar alabilmesidir. Ancak bunun gerçekleşebilmesi için, doğru ve güvenilir bir model geliştirebilmek amacıyla çok büyük miktarda veriye ihtiyaç vardır (Ali vd., 2023). İlk dönemlerde kullanılan makine öğrenmesi yöntemleri — örneğin doğrusal regresyon, lojistik regresyon ve karar ağaçları — verinin düzgün doğrusal veya alt-doğrusal olduğu varsayımına

dayandığı için karmaşık problemler karşısında yeterince başarılı olamıyordu. Oysa gerçek dünya verisi çoğunlukla karmaşık ve doğrusal olmayan bir yapıya sahiptir. Bu nedenle, bu veriden anlamlı bilgi çıkarmak zorlu bir iştir. Tam da bu noktada, karmaşık veri kümelerinden bilgi çıkarabilen derin sinir ağları devreye girer (Saxe, Nelli ve Summerfield, 2021).

Derin sinir ağlarının kullanılmasıyla birlikte, araştırmacılar ağ yapısı derinleştikçe karar verme performansının arttığını fark etmiştir (Piccialli vd., 2021). Ancak bu tür karmaşık veriden anlamlı örüntüler çıkarabilmek için, modelin büyük veri kümeleri üzerinde eğitilmesi gerekir. Gerçek dünya verisinin doğrusal olmayan yapısındaki tüm çeşitliliği yakalayabilmek amacıyla çok sayıda evrişim filtresi ve çekirdek kullanılır; bu da yüksek performanslı yapay zekâ modellerinin ortaya çıkmasını sağlar. Fakat filtre sayısının artması, ağın sonraki katmanlarına daha fazla yük bindirir. Dolayısıyla, temel bir ağ bile çok sayıda katmana, filtreye ve nöron birimine sahip olabilir. Karmaşık görevler için geliştirilen derin sinir ağları ise milyonlarca hatta milyarlarca parametre içerebilir. Ağın katmanları arasındaki veri akışı ve temsil biçimlerini anlamak zordur; üstelik modelin karmaşıklığı arttıkça öğrenilecek değişken sayısı da artar (Li vd., 2021).

Bir derin sinir ağının yapısal tasarımı; kullanılan aktivasyon fonksiyonları, giriş verisinin türü ve boyutu, katman sayısı, havuzlama yöntemleri, bağlantı yapısı, sınıflandırma mekanizmaları ve karma öğrenme teknikleri gibi birçok etkene bağlıdır. Ayrıca, normalleştirme ve düzenleme yöntemleri, ağırlık güncelleme mekanizmaları, maliyet veya kayıp fonksiyonları ve kullanılan son sınıflandırıcı türü de öğrenme sürecini şekillendirir. Tüm bu karmaşık yapı, karar ağaçları, bulanık kural tabanlı sistemler veya Bayesçi ağlar gibi daha anlaşılır yöntemlere kıyasla, derin sinir ağlarının verdiği kararların nedenini anlamayı zorlaştırır. İşte bu durum, literatürde “kara kutu” sorunu olarak bilinir (Vilone ve Longo, 2021).

Öte yandan, karar ağaçları gibi daha basit makine öğrenmesi yöntemlerini anlamak çok daha kolaydır. Burada “anlamak”, modelin aldığı kararları mantıklı ve makul bir şekilde açıklayabilmek anlamına gelir (Doshi-Velez ve Kim, 2017). Kara kutu modellerin aksine, basit modellerin çalışma prensipleri, modelin özeti ya da parametreleri incelenerek kolayca anlaşılabilir; ek bir açıklama modeline ihtiyaç duyulmaz.

Yapay zekâ modellerinin kara kutu doğasından kaynaklanan yorumlanabilirlik eksikliğini gidermek amacıyla, çalışmada yalnızca derin öğrenme tabanlı sınıflandırma sonuçlarına bağlı kalınmamış; bunun yanı sıra klinik değerlendirmeyi destekleyici alternatif bir karar yapısı da geliştirilmiştir. Bu amaçla, konvolüsyonel sinir ağı ve kapılı tekrarlayan ünite (CNN-GRU) mimarisine sahip model tarafından gerçekleştirilen nistagmus sınıflandırmasına ek olarak, bulanık mantık tabanlı bir sistem entegre edilmiştir. Söz konusu sistemde, SAM2 sonucu elde edilen gözbebeği maskelerinin merkez noktaları referans alınarak yatay ve dikey eksenlerde hesaplanan yavaş faz hızı (Slow Phase Velocity, SPV) değerleri bulanık mantık sistemine giriş değişkeni olarak tanımlanmıştır. SPV değerleri için oluşturulan üyelik fonksiyonları ve tanımlanan bulanık kurallar aracılığıyla, hareketin yönü ve şiddetine ilişkin yorumlanabilir kararlar üretilmiştir. Böylece klinisyenlerin sistem tarafından üretilen sayısal ve grafiksel çıktıları doğrudan gözlemleyerek kendi uzman değerlendirmeleriyle ilişkilendirmeleri mümkün hâle getirilmiştir.

4.5.1. Yavaş Faz Hızı (Slow Phase Velocity - SPV)

Gözlerin, hareketli bir hedefi veya sahneyi takip ederken düzgün bir şekilde hareket ettiği döneme SPV denir. SPV, vestibüler sistemin bir uyarana karşısında göz hareketlerini nasıl kontrol ettiğini gösterir. BPPV'nin hangi yarım daire kanalında bulunduğu belirlenmesine yardımcı olur.

Vestibüler fonksiyonların doğru şekilde yorumlanabilmesi için nistagmusun şiddeti, yani yavaş faz hızı önemlidir. SPV, gözlerin yavaş fazında ne kadar hızla hareket ettiğini ifade eder. Nistagmus sırasında gözlerin yavaş hareket ettiği dönemdeki hızıdır. Bu, vestibüler sistemdeki anormal uyarılara göre belirlenir. SPV'yi ölçmek için, gözlerin yavaş hareket ettiği dönem boyunca aldığı yolun, bu hareketin süresine bölünmesi gerekir. Bu ölçüm, genellikle bir saniyelik bir zaman dilimi içinde göz hareketini değerlendirerek yapılır.

$$SPV = \frac{\Delta x}{\Delta t} \quad (4.14)$$

Nistagmusun yavaş fazı, gözlerin nispeten daha yavaş hareket ettiği dönemi ifade eder. Grafikte yavaş faz sırasında gözlerin hareket ettiği mesafe (Δx) ve bu hareketin süresi (Δt) belirtilmiştir. Bu ölçümler, yavaş fazın SPV'sini hesaplamak için kullanılır. SPV,

yavaş faz sırasında gözlerin hareket ettiği mesafenin, bu hareketin süresi ($\Delta x / \Delta t$) oranlanmasıyla elde edilir.

Yatay (x) ve dikey (y) bileşenler için SPV hesaplaması, göz bebeği merkezinin ardışık kareler arasındaki konum değişimlerine dayanmaktadır. Bu kapsamda, SAM2 modeliyle elde edilen iris maskeleri kullanılarak her bir karede göz bebeği merkezi tespit edilmiş, x ve y koordinatlarındaki zamana bağlı konum değişimleri hesaplanmıştır. Elde edilen konum serileri, türev alma işlemiyle hız bileşenlerine (piksel/saniye) dönüştürülmüştür.

Torsiyonel bileşen için SPV hesaplaması, gözün kendi eksenini etrafındaki dönme hızının optik akış yöntemiyle belirlenmesine dayanmaktadır. Bu kapsamda, iris maskesi üzerinden en büyük kontura elips uydurularak göz merkezinin koordinatları ve yaklaşık yarıçap değeri elde edilmiştir. Ardışık kareler arasında Farneback optik akış algoritması uygulanmış, böylece iris çevresindeki piksel hareketlerinin yön ve büyüklük bilgileri çıkarılmıştır. Yatay ve dikey hareket analizlerinde doğrudan SAM2 tarafından üretilen maske görüntüleri kullanılmışken, torsiyonel analizde iris dokusunun yüzey desenlerinin daha belirgin hale gelmesi amacıyla SAM2 maskesi ile orijinal görüntü birleştirilmiş, bu sayede optik akışın dokusal farklılıklara karşı daha duyarlı olması sağlanmıştır. Gözün merkezine göre dairesel doğrultuda gerçekleşen bu hareketler, gözün dönme hareketini temsil eden açısal hızın hesaplanmasında kullanılmıştır. Elde edilen ortalama dönme hızı derece cinsine çevrilip zaman adımına (Δt) bölünerek SPV-T ($^{\circ}/s$) değeri hesaplanmıştır.

Yatay, dikey ve torsiyonel bileşenlerde elde edilen hız sinyalleri, analiz öncesinde Savitzky–Golay filtresi uygulanarak yumuşatılmış ve yüksek frekanslı gürültü bileşenlerinden arındırılmıştır. Ardından, her biri 5 saniyelik zaman pencereleri oluşturulmuş; her pencere içerisinde mutlak hız değerlerinin 75. yüzdeleri dilimi hesaplanarak, nistagmus hareketinin ani dalgalanmalarından etkilenmeyen, daha kararlı bir ortalama hız ölçütü elde edilmiştir. Bu istatistiksel değerler, yatay (SPV-X), dikey (SPV-Y) ve torsiyonel (SPV-T) bileşenler için ayrı ayrı belirlenmiş ve bulanık mantık sistemine nistagmusun şiddetini temsil eden giriş değişkenleri olarak aktarılmıştır.

4.5.2. Bulanık Mantık

Derin öğrenme tabanlı yöntemler, özellikle karmaşık görsel verilerin sınıflandırılmasında yüksek başarı sağlamaktadır. Ancak bu yöntemlerin karar süreci çoğunlukla yorumlanamaz nitelikte olduğundan, özellikle klinik pratikte doğrudan kullanımları bazı çekincelere yol açmaktadır. Klinik çalışanlar, yapay zekâ tabanlı sistemlerin verdiği kararları doğrudan uygulamak yerine, süreci anlamlandırmak ve kendi klinik değerlendirmeleri ile desteklemek isteyebilir. Bu bağlamda, bu çalışmada önerilen bulanık mantık tabanlı yaklaşım, hem sayısal hem de grafiksel olarak anlamlı çıktılar üretmekte ve hekimlerin karar sürecine aktif katılımını desteklemektedir. Ortalama SPV değerlerine dayalı olarak oluşturulan puanlama sistemi, bulanık kurallar aracılığıyla yorumlanabilir bir karar yapısı sunmakta ve derin öğrenme tabanlı sınıflandırma sonuçlarını destekleyici nitelikte alternatif bir değerlendirme mekanizması oluşturmaktadır.

Bulanık mantık, klasik mantık sistemlerinin kesinlik üzerine kurulu yapısından farklı olarak, belirsizlikleri ve ara değerleri de dikkate alan esnek bir akıl yürütme yöntemidir (Saatchi, 2024). Zihinsel modellemesini insan düşünce biçiminden alan bu yaklaşım, gerçek dünyada sıklıkla karşılaşılan “kısmen doğru” ya da “yaklaşık” gibi niteliksel ifadelerin sayısal sistemlere entegre edilebilmesini sağlar. Klasik mantıkta bir ifade yalnızca “doğru” (1) ya da “yanlış” (0) olarak değerlendirilebilirken, bulanık mantıkta bir durum 0 ile 1 arasında herhangi bir değeri alabilir. Böylece, gerçek hayattaki belirsizlik içeren ya da net sınırlarla tanımlanamayan durumlar daha etkin bir şekilde modellenebilir.

Bulanık mantık kavramı ilk olarak 1965 yılında Lotfi A. Zadeh tarafından ortaya atılmıştır (Kumar, Raj ve Mohammadzadeh, 2024). Zadeh’in geliştirdiği bu teori, klasik doğruluk sistemlerinin yetersiz kaldığı alanlarda, özellikle de insan sezgisinin ve uzman bilgisinin önemli olduğu sistemlerde yaygın şekilde uygulanmaya başlanmıştır. Günlük hayatta karşılaştığımız birçok kavram, keskin sınırlar yerine derece farkları içerir. Örneğin “sıcak hava” tanımını herkes için farklı anlamlar taşıyabilir; 26 °C bir kişi için “ılık”, bir başkası için “sıcak” olarak değerlendirilebilir. Bu tür yorumlara açık durumlar, klasik mantıkla modellenemezken, bulanık mantık bu tür kademeli sınıflandırmaları kolaylıkla işleyebilir (Saatchi, 2024).

Bulanık mantık sistemlerinin temelinde, belirsizlik taşıyan verilerle rasyonel çıkarımlar yapılabilmesini sağlayan bir yapı bulunmaktadır. Bu yapı genellikle dört temel adımdan oluşur. İlk aşamada, sistemin aldığı sayısal girişler, tanımlanmış üyelik fonksiyonları aracılığıyla uygun bulanık kümelere dönüştürülür. Bu işleme “bulanıklaştırma” (fuzzification) adı verilir. Ardından, sistemin davranışlarını yönlendiren ve uzman bilgisine dayanan “eğer...ise...” şeklindeki kural tabanı devreye girer. Bu kurallar, giriş değişkenlerinin durumuna göre sistemin nasıl tepki vereceğini tanımlar. Üçüncü adımda çıkarım mekanizması devreye girerek uygun kuralları işler ve olası bulanık sonuçları üretir. Son aşamada ise elde edilen bulanık çıktılar, tek bir sayısal değere indirgenir; bu işleme “keskinleştirme” (defuzzification) denir. Böylece sistemin dış dünyaya vereceği tepki netleştirilmiş olur (Kumar, Raj ve Mohammadzadeh, 2024).

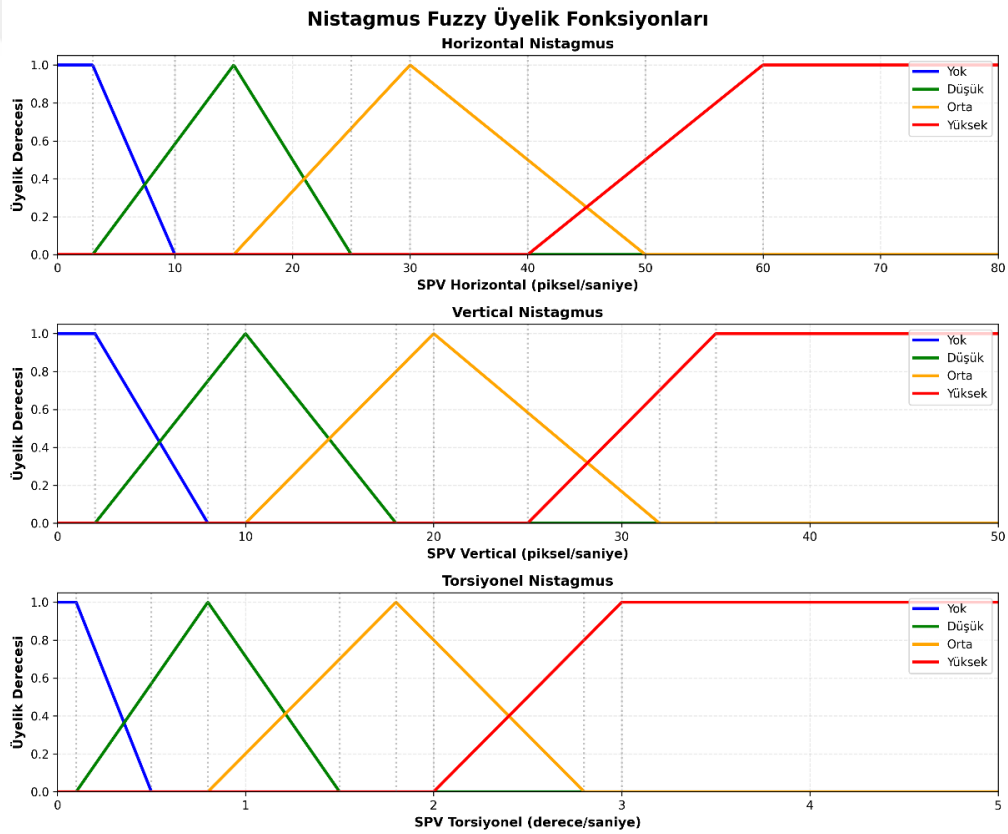
Bulanık mantığın en dikkat çeken yönlerinden biri, karar süreçlerinde insan benzeri sezgisel yaklaşımları modelleyebilmesidir. Bu sayede, doğrusal ya da iyi tanımlanmış matematiksel modellere ihtiyaç duyulmaksızın, karmaşık sistemlerde dahi başarılı sonuçlar elde edilebilir. Özellikle giriş verilerinin kesin olmadığı veya sistemin davranışı klasik yöntemlerle modellenemeyecek kadar karmaşık olduğu durumlarda, bulanık mantık güçlü bir çözüm yöntemi olarak öne çıkmaktadır.

Bu yöntem, mühendislikten yapay zekâya, tıbbi teşhis sistemlerinden otomatik kontrol uygulamalarına kadar geniş bir yelpazede uygulanmaktadır (Castillo vd., 2024). Örneğin, bir klima sisteminde ortam sıcaklığına bağlı olarak fan hızının ayarlanması, bir çamaşır makinesinde yıkama süresinin belirlenmesi ya da bir robotun insan davranışlarını taklit edebilmesi gibi uygulamalar, bulanık mantık kuralları ile oldukça verimli şekilde gerçekleştirilebilir. Bu tür sistemlerde, uzmanlardan alınan bilgi doğrudan kurallara dönüştürülerek sisteme entegre edilebilir. Bu durum, hem sistemin açıklanabilirliğini artırır hem de geliştirme sürecinde sezgisel bir yapı sunar.

Ancak bulanık mantığın başarıya ulaşabilmesi, büyük ölçüde kullanılan üyelik fonksiyonlarının doğru tanımlanmasına ve kural tabanının etkinliğine bağlıdır. Bu nedenle sistemin performansı, uzman bilgisiyle oluşturulan mantık kurallarının doğruluğuna ve verilerin uygun şekilde yorumlanmasına doğrudan bağlıdır. Yine de bulanık mantığın avantajları, özellikle keskin sınırlarla tanımlanamayan problemlerde

sağladığı çözüm kapasitesi sayesinde, onu modern yapay zekâ ve kontrol sistemlerinde vazgeçilmez bir araç haline getirmiştir (Castillo vd., 2024).

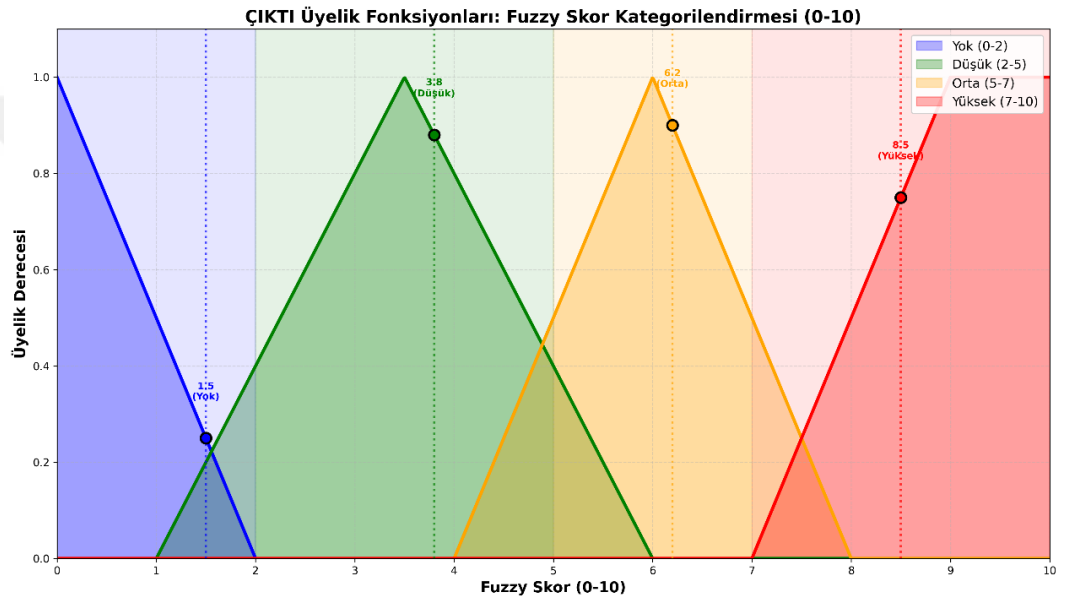
Bu çalışmada geliştirilen bulanık mantık modülü, üç giriş ve bir çıkış değişkeninden oluşmaktadır. Giriş değişkenleri, nistagmus hareketinin her bir eksenine karşılık gelen SPV-X, SPV-Y ve SPV-T (torsiyonel SPV) değerleridir. Her bir SPV değişkeni sırasıyla yatay, dikey ve torsiyonel eksenlerdeki yavaş faz hızını temsil etmekte olup, 0–60 px/s aralığında sürekli değerler alır (torsiyonel eksen için birim derece/saniye cinsindedir). Bu giriş değişkenleri, her 5 saniyelik zaman penceresi sonunda elde edilen ortalama SPV değerleri üzerinden tanımlanan üyelik fonksiyonları aracılığıyla değerlendirilir. Her eksene ilişkin bulanık üyelik dereceleri (“yok”, “düşük”, “orta”, “yüksek”) ayrı ayrı hesaplanarak, kural tabanlı çıkarım mekanizmasıyla bütünleştirilir.



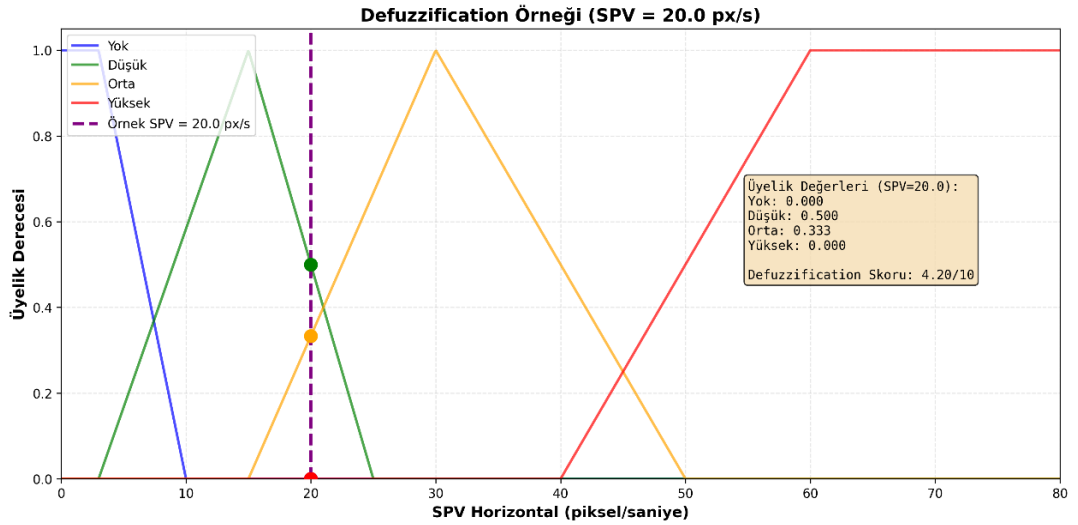
Şekil 4.12: Bulanık mantık hareket bazlı üyelik fonksiyonları.

Bu çalışmada kullanılan üyelik fonksiyonları, **üçgensel (triangular)** ve **omuz (shoulder)** tipindedir. Her bir eksen için dört üyelik kümesi tanımlanmıştır: “yok”, “düşük”, “orta” ve “yüksek”. Yatay eksen (SPV-X) için bu fonksiyonlar sırasıyla [3, 10], [3, 15, 25], [15, 30, 50] ve [40, 60]; dikey eksen (SPV-Y) için [2, 8], [2, 10, 18],

[10, 20, 32] ve [25, 35]; torsiyonel eksen (SPV-T) için ise derece/saniye cinsinden [0.1, 0.5], [0.1, 0.8, 1.5], [0.8, 1.8, 2.8] ve [2.0, 3.0] parametreleriyle modellenmiştir. Bu değerler, klinik literatürde bildirilen hareket aralıklarına uygun olacak şekilde seçilmiştir. Her eksene ait SPV değeri kendi üyelik fonksiyonları üzerinden değerlendirilerek dört üyelik derecesi hesaplanır. Elde edilen üyelik dereceleri **ağırlıklı ortalama (weighted average)** yöntemiyle durulaştırılır (defuzzification) ve 0–10 aralığında tek bir sayısal skor (fuzzy skor) üretilir. Ağırlık değerleri sırasıyla “yok = 0”, “düşük = 3”, “orta = 6” ve “yüksek = 9” olarak atanmıştır. Böylece sistem, her ekseninde ölçülen SPV değerini klinik açıdan anlamlı bir skora dönüştürür.



Şekil 4.13: Çıktı üyelik fonksiyonları ve fuzzy skor kategorilendirmesi (0-10).



Şekil 4.14: Üyelik derecesine göre defüzyon örneği.

Şekil 4.13 ve Şekil 4.14’te, yatay, dikey ve torsiyonel eksenlere ait üyelik fonksiyonlarının parametreleşmesi ile etkin SPV değerlerinin bu üyelik derecelerini nasıl etkilediği ve ağırlıklı ortalama yöntemi kullanılarak 0–10 aralığında tek bir skora dönüştürüldüğü, örnek bir senaryo üzerinden gösterilmektedir.

Üretilen skor değerleri, klinik olarak anlamlı eşiklerle yorumlanmıştır. Buna göre, 5.5’in altındaki skorlar “düşük şiddette nistagmus”, 5.5–7.5 aralığındaki skorlar “orta şiddette nistagmus”, 7.5’in üzerindeki skorlar ise “yüksek şiddette nistagmus” olarak değerlendirilmiştir. Üç eksen den elde edilen skorlar karşılaştırılarak nistagmusun yönü ve olası karma yapısı belirlenmiştir. Bir eksenin anlamlı sayılabilmesi için skorunun 4.0’ın üzerinde olması gerekmektedir. Baskın eksen dışındaki bileşenlerden herhangi biri, hem 3.0’ın üzerinde skor alıyor hem de baskın eksenin skorunun %50’sinden fazlasına ulaşıyorsa, bu durum karma nistagmus (örneğin horizontal + torsional) olarak sınıflandırılmıştır. Sistem ayrıca baskın eksen bilgisini —örneğin “yüksek şiddette torsiyonel nistagmus” ifadesinde olduğu gibi— klinik yoruma dahil ederek hem sayısal hem de sözel çıktı üretmektedir. Böylece geliştirilen model, nistagmusun şiddetini, yönünü ve bileşenlerini açıklanabilir biçimde sunarak klinik karar sürecini destekleyen güvenilir bir araç hâline gelmiştir.

BEŞİNCİ BÖLÜM

DENEYSEL SONUÇLAR VE ANALİZLER

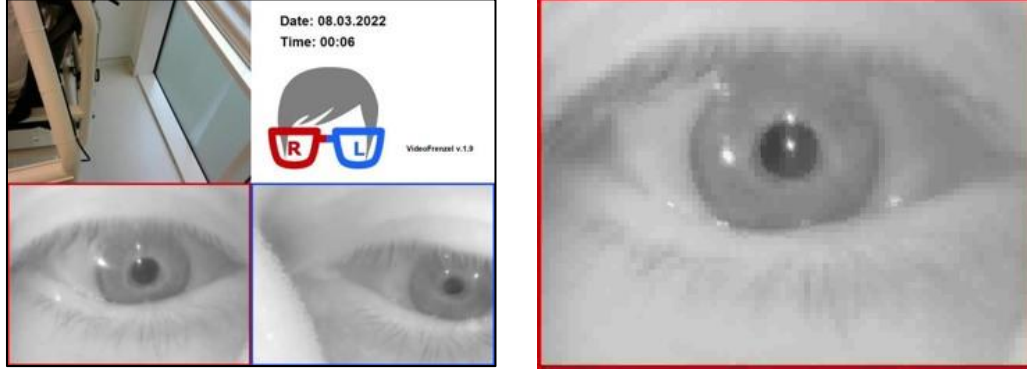
Bu bölümde, çalışmada kullanılan BPPV veri kümesi tanıtılmakta, uygulanan göz segmentasyonu yöntemlerinin performansı değerlendirilmektedir ve optik akış analizinin sonuçları tartışılmaktadır. İlk olarak, BPPV tanısı için kullanılan video verisinin içeriği ve ön işleme adımları açıklanmaktadır. Ardından, SAM2 modeli kullanılarak elde edilen segmentasyon sonuçları görsel olarak değerlendirilmiş ve özellikle zorlu klinik senaryolar altındaki performansına odaklanılmıştır.

Son olarak, optik akış yöntemiyle elde edilen hareket vektörleri analiz edilmiş ve nistagmus örüntüleri HSV renk uzayında görselleştirilmiştir. Buna ek olarak, elde edilen optik akış öznitelik haritaları, CNN-GRU tabanlı bir sınıflandırma modeli kullanılarak dört farklı nistagmus yönünü sınıflandırmak amacıyla kullanılmıştır.

5.1. Veri seti ve Ön işleme

Nistagmus analizi, BPPV'nin tanı sürecinde doğru şekilde sınıflandırılması için kritik öneme sahiptir (You, Instrum ve Parnes, 2019). Hastalara kontrollü baş manevraları uygulanarak göz hareketleri gözlemlenir ve etkilenen vestibüler kanalın—arka, yatay veya ön kanal—hangisi olduğu belirlenmeye çalışılır. BPPV, iç kulaktaki bu üç yarım daire kanalından herhangi birinde ortaya çıkabileceğinden, belirli pozisyonel manevralara verilen oküler tepkilerin analiz edilmesi, kesin tanı için hayati önem taşır.

Bu çalışmada, etiketlenmiş bir video veri kümesi kullanılmıştır. Bu veri kümesi, BPPV'nin hem tanısını hem de tedavisini tek bir seansta gerçekleştirmeyi amaçlayan, yapay zekâ destekli ve hassasiyetle kontrol edilen bir platform olan Robotic Maneuvering System (RMS) kullanılarak elde edilmiştir (Bayram vd., 2025; Özkul, 2023).



(a)

(b)

Şekil 5.1: Veri kümesinden ham görüntü karesi: (a) kırpılmış göz bölgesi: (b).

Göz hareketleri, özel tasarlanmış gözlükler kullanılarak RGB formatında ve 640×480 piksel çözünürlükte kaydedilmiştir. Her bir video, uygulanan manevraya ve göz hareketlerinin yönüne göre etiketlenmiş olup, etkilenen kanalla ilişkili yapılandırılmış bir sınıflandırma sürecine olanak tanımaktadır. Kaydedilen verilerden örnek bir kare ve buna karşılık gelen kırpılmış göz bölgesi, Şekil 5.1’de gösterilmiştir.

Ham video kayıtları, tanı sürecinin çeşitli aşamalarını içermektedir; bu aşamalar arasında hastanın sandalyeye oturtulması, gözlüklerin takılması, hekim–hasta iletişimi ve BPPV tanısında kullanılan üç temel manevranın uygulanması yer almaktadır. Bu manevralar şunlardır:

- Dix–Hallpike (arka kanal),
- Supine Head-Roll (yatay kanal),
- Deep Head Hanging (ön kanal).

Her bir manevra, belirli bir süre boyunca uygulanmış ve bu esnada hastanın göz hareketleri kesintisiz olarak kaydedilmiştir. Ancak, ham görüntüler tanıya yönelik olmayan bölümler de içerdiğinden, analizden önce bir ön işleme (preprocessing) aşaması gereklidir. Bu aşamada, yalnızca gerçek manevralara karşılık gelen bölümler çıkarılmış ve her test için yaklaşık birer dakikalık video parçaları elde edilmiştir. Her manevranın başlangıç ve bitiş noktaları manuel olarak belirlenmiştir. Tutarlılığı sağlamak amacıyla, her hastadan yalnızca bir göz seçilerek analiz yapılmıştır; çünkü nistagmus, genellikle her iki gözde senkronize olarak gözlemlenir.

Tablo 5.1: BPPV veri kümesinde her bir nistagmus sınıfına ait örnek sayısı.

Sınıf Etiketi	Örnek Sayısı
Horizontal	349
Vertical	241
Torsional	381
Vertical + Torsional	264

Sınıflandırma sürecinde, her bir 1 dakikalık video 300 kareye bölünerek modelin giriş gereksinimlerine uygun hale getirilmiştir. Bu kareleme işlemi, her sınıf için elde edilen örnek sayısının temelini oluşturmaktadır. Sınıf başına örnek dağılımının özeti, Tablo 5.1’de sunulmuştur.

Ön işleme sonrasında, göz bebeği segmentasyon yöntemlerinin gerçekçi klinik koşullar altındaki dayanıklılığını değerlendirmek amacıyla beş farklı zorlu senaryo tanımlanmıştır. Bu senaryolar şunlardır:

- **Şiddetli Nistagmus Hareketi Senaryosu:** Hızlı ve geniş açılı nistagmus ile karakterize edilir; bu durum, göz segmentasyonunun kararlılığını ve hassasiyetini önemli ölçüde zorlaştırır.
- **Düşük Göz Kapağı ve Sınırlı Görünürlük Senaryosu:** Düşük göz kapağı konumu göz bebeğini kısmen örttüğünde ortaya çıkar; bu durum görünürlüğü azaltır ve segmentasyon doğruluğunu karmaşık hale getirir.
- **Kısmen Kapalı Göz Kapağı ile Şiddetli Nistagmus Senaryosu:** Hızlı nistagmus ve kısmen kapalı göz kapağının eşzamanlı olarak görülmesiyle oluşur; bu birleşim, görsel engel ve hareket bulanıklığını bir arada oluşturur.
- **Yoğun Kirpik ve Sınırlı Göz Açıklığı Senaryosu:** Yoğun kirpikler ve dar bir göz aralığı, gölgelenme ve fiziksel örtülme oluşturduğunda ortaya çıkar; bu da modelin göz bebeği veya iris sınırını doğru şekilde tespit etmesini engeller.
- **Makyaja Bağlı Zorluk Senaryosu:** Göz makyajının, göz çevresinde ani kontrast değişimlerine yol açtığı koşulları yansıtır; bu durum, segmentasyon algoritmalarını yanıltabilir.

Klinik ortamlarda, göz kapağının kısmen kapalı olması ya da kirpiklerin göz bebeğini örtmesi gibi çevresel faktörler, göz bebeği görünürlüğünü azaltarak manuel değerlendirmeyi zorlaştırır. Bu durum, otomatik segmentasyon modellerinin de hatalı

sonular retmesine neden olabilir ve potansiyel olarak hastanın yanlış tanı almasına yol aabilir. rneęin, gz kresi hareketi yerine gz kapaęı hareketinin tespit edilmesi, nistagmus analizinde hatalara neden olabilir.

Őiddetli nistagmus durumunda gz bebeęinin ani ve dzensiz hareketleri, segmentasyon algoritmalarının kararlılıęını daha da zorlaŐtırır. Hatalı segmentasyon, gz bebeęinin yanlış konumlandırılmasına neden olabilir ve bu da hareket yn analizini yanıltabilir. Bu zorlukların stesinden gelmek amacıyla, hem YOLO tabanlı modeller hem de SAM2 modeli tanımlanan zorlu senaryolar altında test edilerek dayanıklılıkları deęerlendirilmiŐtir.

YOLO modelleri, mimari farklılıklarına gre deęerlendirilmiŐ ve hız ile segmentasyon doęruluęu aısından karŐılaŐtırılmıŐtır. Olumsuz koŐullar altındaki performansın daha ayrıntılı incelenmesi amacıyla, SAM2 ek bir segmentasyon yntemi olarak dhil edilmiŐtir. BeŐ senaryo boyunca her iki yaklaŐımın karŐılaŐtırmalı analizi gerekleŐtirilmiŐ ve model dayanıklılıęı ayrıntılı biimde deęerlendirilmiŐtir.

Segmentasyon adımıını takiben, her modelin ıktıları kullanılarak optik akıŐ analizi gerekleŐtirilmiŐ ve ortaya ıkan hareket znelik haritaları zerinde drt sınıflı nistagmus sınıflandırması yapılmıŐtır. Bu grevde kullanılan CNN-GRU mimarisi, YOLOv8, YOLOv11, YOLOv12 ve SAM2 tarafından retilen segmentasyonların sınıflandırma performansını deęerlendirmek ve karŐılaŐtırmak amacıyla kullanılmıŐtır.

5.2. YOLO Segmentasyon Yntemlerinin KarŐılaŐtırılması

Bu blmde, aynı veri kmesi zerinde eęitilen  farklı YOLO segmentasyon modelinin (YOLOv8-seg, YOLOv11-seg ve YOLOv12-seg) performansları karŐılaŐtırılmıŐtır. KarŐılaŐtırma; eęitim srecinde elde edilen kayıp deęerleri (loss), segmentasyon doęruluęu (mAP50-95) ve ıkarım sresi gibi ltler dikkate alınarak yapılmıŐtır. Ama, segmentasyon baŐarımı ile hız arasındaki dengeyi en iyi saęlayan modeli belirleyerek, belirli BPPV'ye ynelik kullanım senaryoları iin etkili segmentasyon ynteminin seilmesine katkı sunmaktır.

alıŐmada, gzbebeęi tespiti amacıyla segmentasyon modelleri, CASIA-Iris-Interval veri seti kullanılarak eęitilmiŐtir. Her  model de aynı eęitim ve doęrulama alt kmeleri zerinde, 50 epoch boyunca aynı hiperparametre ayarlarıyla eęitilmiŐ ve deęerlendirilmiŐtir. Bylece modellerin eęitimi sırasında ortaya ıkan farkların yalnızca mimari yapıdan ve ęrenme yeteneklerinden kaynaklanması saęlanmıŐtır.

Eğitim süreci sonunda, modeller segmentasyon başarımı, model karmaşıklığı ve latency gibi çeşitli metrikler üzerinden karşılaştırılmıştır. Aşağıda sunulan tablo, bu karşılaştırmanın genel bir özetini sunmakta ve farklı uygulama senaryoları için hangi modelin daha uygun olabileceğine dair içgörüler sağlamaktadır.

Model performansı karşılaştırılırken, mAP değeri hem nesne algılama (box) hem de segmentasyon (mask) için 50-95 IoU aralığında hesaplanmıştır. Bu çalışmada segmentasyon değerlendirme için mAP@[.50:.95] (mask) metriği tercih edilmiştir. Bu metrik, IoU değerinin 0.5 ile 0.95 arasındaki çeşitli eşiklerde hesaplanmasıyla, farklı hassasiyet seviyelerinde model performansını daha ayrıntılı biçimde ortaya koymaktadır.

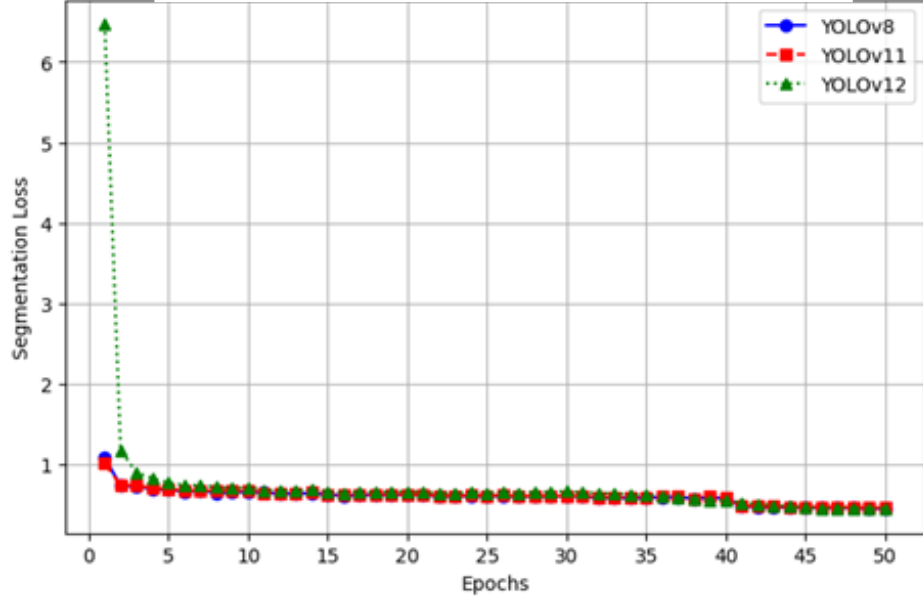
Nesne tabanlı segmentasyon gerçekleştiren YOLO mimarisiyle uyumlu olması nedeniyle, klasik piksel temelli metrikler yerine mAP kullanılması daha anlamlı bir değerlendirme sunmaktadır. Eğitim sonrası hız değerlendirmesi, A100 GPU üzerinde elde edilen gecikme (latency) değerleri ile yapılmıştır. Model karmaşıklığını yansıtan parametre (Params) ve FLOPs (floating point operations) değerleri de tabloya dahil edilerek, modellerin hem doğruluk hem de verimlilik yönünden karşılaştırılmasına olanak sağlanmıştır.

Tablo 5.2: YOLO segmentasyon modellerine ait performans metrikleri.

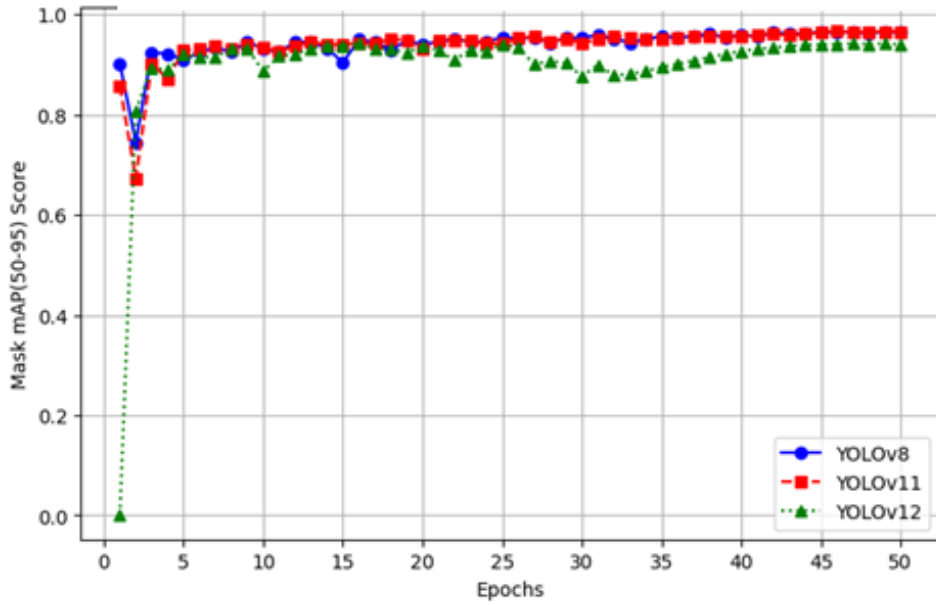
Model	Boyut (piksel)	mAP _{box} (50-95)			mAP _{mask} (50-95)			Hız (ms) A100 GPU	Params Flop's Latency		
		Iris	Pupil	Ort.	Iris	Pupil	Ort.		(M)	(B)	(ms/img)
Yolov8	480x640	0.989	0.974	0.981	0.960	0.971	0.965	1.21x10 ⁶	11.8	42.4	2.3
Yolov11	480x640	0.989	0.975	0.982	0.961	0.974	0.967	1.33x10 ⁶	10.1	35.3	2.3
Yolov12	480x640	0.978	0.962	0.970	0.927	0.955	0.941	1.62x10 ⁶	9.8	35.2	2.8

Tablo 5.2'de, üç farklı YOLO segmentasyon modeline ait karşılaştırmalı metrik sonuçları sunulmuştur. mAP_{mask}(50-95) açısından en yüksek doğruluk değerine YOLOv11-seg modeli ulaşmıştır. YOLOv8-seg ve YOLOv11-seg modelleri 2.3 ms/img gecikme süresi ile benzer bir hız sergilerken, YOLOv12-seg modeli 2.8 ms/img ile daha yavaş bir performans göstermiştir.

Hesaplama yükü (FLOPs) ve parametre sayısı açısından ise YOLOv12-seg, diğer modellere kıyasla daha küçük boyutlu ve hafif bir yapıdadır. Model performansını daha iyi değerlendirebilmek adına, her bir modelin eğitim sürecinde elde edilen segmentasyon başarımları ve kayıp değerleri grafiksel olarak analiz edilmiştir.



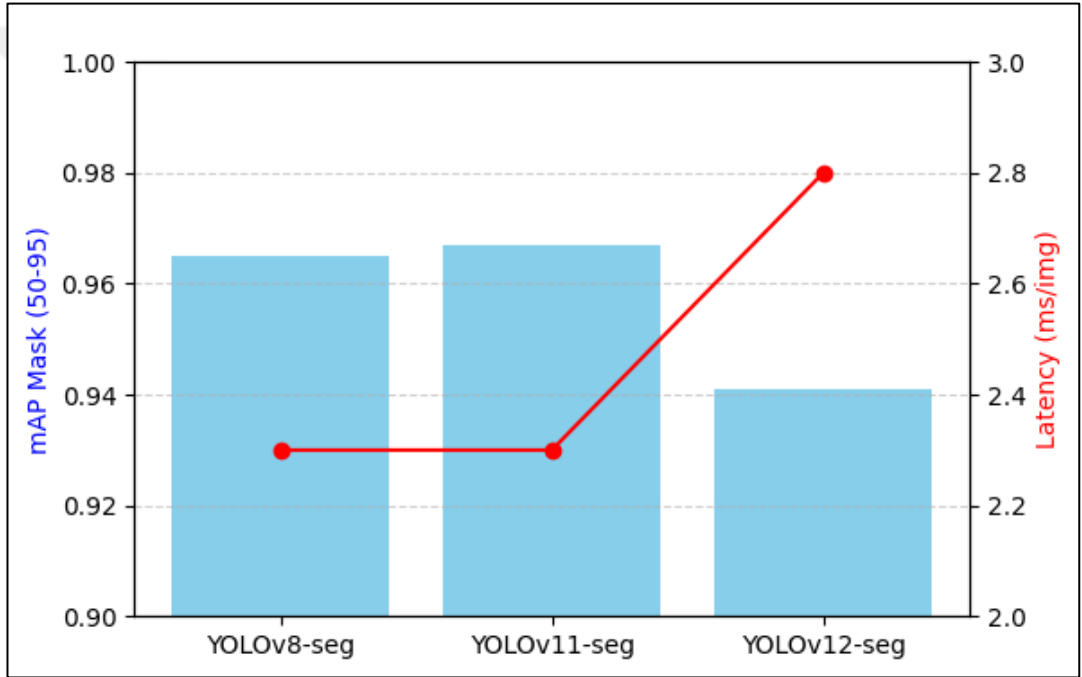
Şekil 5.2: YOLO modellerinin 50 epoch boyunca CASIA-Iris-Interval veri kümesi üzerindeki segmentasyon kayıp (segmentation loss) grafiği.



Şekil 5.3: YOLO modellerinin 50 epoch boyunca CASIA-Iris-Interval veri kümesi üzerindeki maske mAP (50–95) skor grafiği.

Şekil 5.2 ve Şekil 5.3’de, modellerin epoch boyunca Mask mAP(50-95) performansı gösterilmektedir. YOLOv11-seg, eğitimin erken safhalarından itibaren istikrarlı bir şekilde yüksek doğruluk değerlerine ulaşırken, YOLOv12-seg’in daha dalgalı bir öğrenme eğrisi sergilediği gözlenmektedir. Şekil Y’de yer alan seg_loss grafiğinde ise üç modelin de kayıp değerleri zamanla azalarak stabil bir seviyeye geldiği, YOLOv11-seg’in daha hızlı bir şekilde daha düşük loss seviyelerine ulaştığı görülmektedir.

Modellerin sadece doğruluk açısından değil, aynı zamanda tahmin süresi gibi hızla ilgili metrikler açısından da değerlendirilmesi önemlidir. Bu nedenle, Şekil 5.4’te YOLO segmentasyon modellerinin mAP Mask (50–95) skorları ile latency (ms/img) değerleri birlikte görselleştirilmiştir.



Şekil 5.4: YOLO modellerinin doğruluk ve gecikme karşılaştırması.

Grafikte, mAP Mask (50–95) skorları mavi çubuklarla, karşılık gelen gecikme (latency) değerleri (ms/görüntü) ise kırmızı çizgiyle gösterilmiştir. YOLOv11-seg modeli hem en yüksek doğruluk değerine (mAPmask=0.967) hem de düşük gecikme süresine (2.3 ms/img) sahiptir. YOLOv8-seg modeli de benzer doğruluk (0.965) ve gecikme süresi sunarken, YOLOv12-seg modeli daha düşük doğrulukla (0.941) birlikte daha yüksek bir gecikme süresi (2.8 ms/img) göstermektedir.

Bu analiz sonucunda, doğruluk ve işlem süresi açısından en iyi dengeyi sağlayan modelin YOLOv11-seg olduğu görülmektedir. Modelin yüksek doğruluk oranını

korurken daha düşük gecikme süresi sunması, özellikle gözbebeği segmentasyonu gibi zaman hassasiyeti gerektiren BPPV tanı süreçlerinde etkili bir alternatif olabileceğini göstermektedir. Bu nedenle, YOLOv11-seg modeli, hem hassas segmentasyon gereksinimlerini karşılamak hem de işlem süresini minimize etmek isteyen uygulamalar için avantajlı bir seçenek olarak değerlendirilebilir.

Tablo 5.3: CASIA-Iris-Interval veri kümesini kullanan çalışmalar arasında segmentasyon performansı karşılaştırması; CAD: Circular Active Discs, CHT: Circular Hough Transform, ID: Integro-Differential Operator, F1: F1 Skoru.

Referans#	Yöntem	Metrik	Segmentasyon Skoru	
			Iris	Pupil
(Farmanifard ve Ross, 2024)	SAM	IoU	0.969	-
(Mathias vd., 2022)	B-spline + aktif kontur	Doğruluk, Dice	0.946, 0.898	0.994, 0.95
(Balasubramanian vd., 2023)	Özel CNN	Doğruluk, F1	0.992, 0.983	-
(Kumar ve Teotia, 2019)	CAD	Doğruluk, Dice	-	0.947, 0.956
(Jan vd., 2021)	CHT (pupil) + ID (iris)	Doğruluk	0.986	-
(Al-Waisy vd., 2018)	CHT + Daugman Mapping	Doğruluk	0.998	-
Önerilen Çalışma (2025)	Yolov8-seg		0.960	0.971
	Yolov11-seg	mAPmask (50-95)	0.961	0.974
	Yolov12-seg		0.927	0.955

Tablo 5.3’de, CASIA-Iris-Interval kullanılarak göz bölgesi segmentasyonu yapılan çalışmalar karşılaştırılmıştır. Segmentasyon hedefleri, kaynak çalışmaların sunumuna uygun şekilde “Iris”, “Pupil” olarak ayrı ayrı belirtilmiştir. “Iris” ve “Pupil” her bir bölgenin ayrı ayrı değerlendirildiği çalışmaları temsil etmektedir.

Çalışmalarda kullanılan metrikler (Dice, IoU, Doğruluk, F1, mAPmask) farklılık gösterdiğinden, sonuçlar kaynaklarda bildirildiği şekliyle aktarılmış; metrikler arasında doğrudan bir karşılaştırma veya normalizasyon yapılmamıştır. Bu metriklerin

farklılık göstermesi doğrudan nicel karşılaştırmayı sınırlasa da, literatürdeki yöntem çeşitliliğini ortaya koymaktadır. Bu çalışmada kullanılan mAP@[.50:.95] (mask) metriği, YOLO tabanlı segmentasyon modelinin nesne odaklı yapısına daha uygun bir değerlendirme sunmakta ve farklı IoU eşiklerinde ayrıntılı bir performans analizi sağlamaktadır.

5.3. YOLO ve SAM2 Segmentasyon Performansının Değerlendirilmesi








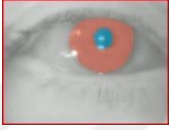


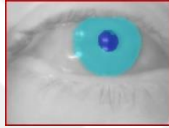

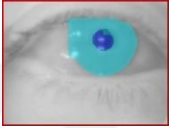


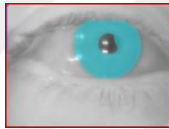
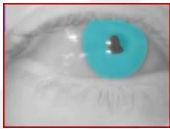
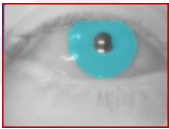

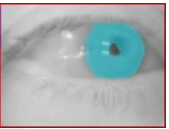





Segmentasyon, BPPV tanısında nistagmus hareketlerinin hassas bir şekilde analiz edilmesi için temel bir adımdır. Bu bölümde, göz bebeği segmentasyonu için SAM2 ve YOLO tabanlı segmentasyon modelleri kullanılmış ve elde edilen sonuçlar görsel olarak değerlendirilmiştir. SAM2, detaylı bölütleme sağlarken, YOLO hızlı ve verimli bir segmentasyon süreci sunmaktadır.

Segmentasyon işleminin ardından, göz bebeğinin hareket özelliklerini çıkarmak amacıyla optik akış analizi uygulanmıştır. Optik akıştan elde edilen hareket vektörleri, HSV renk uzayında temsil edilerek göz bebeğinin zamansal ve mekânsal hareket paternleri görselleştirilmiştir. YOLO ve SAM2 tabanlı segmentasyon sonuçları, farklı senaryolar altında incelenmiş ve bu yöntemlerin doğruluk, hız ve dayanıklılık açısından karşılaştırması yapılmıştır.

Bu çalışmada, farklı çevresel koşullarda göz bebeği tespitinin performansını değerlendirmek amacıyla beş farklı senaryo belirlenmiştir. Bu senaryolar, hem literatürde mevcut çalışmaların başarımında sınırlamalar yaşanan durumları hem de uzman görüşleri doğrultusunda klinik tanıda zorluk yaratan koşulları göz önüne alarak seçilmiştir. Segmentasyon modellerinin çevresel faktörlere ve hızlı nistagmus hareketlerine karşı dayanıklılığını test etmek için hazırlanan bu senaryolar, BPPV tanısında kullanılan hasta verilerinden seçilmiş ve farklı görsel engelleri içeren videolar üzerinde uygulanmıştır.

1. Şiddetli Nistagmus Hareketi: Bu senaryoda, göz bebeği net bir şekilde görülebilmekte ancak nistagmus hareketleri oldukça şiddetlidir. Hızlı ve geniş açılı hareketler, segmentasyon modelinin göz bebeğini sürekli takip etmesini zorlaştırmaktadır. Geleneksel yöntemlerde, hareketin hızı nedeniyle göz bebeğinin konumunun yanlış tahmin edilmesi veya kısmen kaybolması gibi sorunlar ortaya çıkmaktadır. Bu durumda, segmentasyon modelinin göz bebeğini doğru bir şekilde takip edip edemediği değerlendirilmektedir.

Görsel çıktılar incelendiğinde, göz bebeğinin belirgin olmasının segmentasyonu kolaylaştırdığı, ancak ani ve geniş açılı hareketlerin bazı karelerde göz bebeği sınırlarında hatalara yol açtığı gözlemlenmiştir. Özellikle nistagmusun en hızlı olduğu anlarda, modellerin segmentasyon maskelerinde küçük kaymalar veya kısmi eksiklikler meydana gelmiştir.

	Kare 1	Kare 2	Kare 3	Kare 4	Kare 5
Orijinal Kare					
SAM2					
YOLOv8					
YOLOv11					
YOLOv12					

Şekil 5.5: Şiddetli nistagmus hareket senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.

SAM2 modeli genel olarak iris ve pupil bölgelerini her karede tutarlı olarak segment edebilmiştir. Bununla birlikte nistagmusun yoğunluğundan kaynaklanan ufak kaymalar görülsede, maskelerde eksik ya da tamamen başarısız bir segmentasyon gözlenmemiştir. YOLO tabanlı modeller ise iris bölgesini çoğunlukla tutarlı şekilde segment etmiş fakat pupil bölgesinde farklı performanslar göstermiştir.

YOLOv8-seg iris segmentasyonunu başarılı şekilde gerçekleştirirken, pupil maskelerinde eksiklikler gözlenmiştir. YOLOv11-seg modeli ise pupil bölgesini hiç segment edemezken, iris bölgesinde segmentasyonu genellikle sağlayabilmiş fakat bazı karelerde belirgin kaymalar oluşmuştur.


























YOLOv12-seg modeli ise pupil segmentasyonunda YOLOv8-seg ve YOLOv11-seg modellerine göre daha başarılı sonuçlar üretmiştir. Ancak YOLOv12-seg'in iris segmentasyonunda daha fazla kayma ve düzensizlik meydana geldiği gözlenmiştir.

Sonuçlar, SAM2'nin hızlı hareketler içeren senaryolarda bile yüksek doğruluk sağladığını ancak aşırı hızlı ve ani hareketlerden kısmen etkilendiğini göstermektedir. YOLO modelleri ise iris segmentasyonunda tutarlı olmakla beraber pupil segmentasyonunda şiddetli nistagmus hareketlerinden ciddi ölçüde etkilenmiştir.

2. Düşük Göz Kapağı ve Kısıtlı Görünürlük: Bu senaryoda, hastanın göz kapağı oldukça düşük olup göz bebeğinin büyük bir kısmını örtmektedir. Düşük göz kapağı, göz bebeğinin görünürlüğünü azaltarak segmentasyon modelinin başarılı bir şekilde çalışmasını zorlaştırmaktadır. Geleneksel yöntemlerde, göz kapağının oluşturduğu gölgeler ve belirsiz sınırlar nedeniyle yanlış segmentasyon sonuçları elde edilebilir.

SAM2'nin performansı incelendiğinde, göz kapağının kısmen kapattığı karelerde modelin göz bebeğini başarılı bir şekilde tespit etmeye devam ettiği görülmektedir. Hatta göz bebeğinin çok küçük bir kısmının görüldüğü durumlarda dahi modelin segmentasyonu sürdürebildiği dikkat çekmektedir. Bu, modelin çevresel faktörlere karşı belirli bir dayanıklılığa sahip olduğunu göstermektedir. Ancak, en uç noktada göz bebeğinin neredeyse tamamen kapanmasıyla segmentasyon alanının küçüldüğü ve doğruluğun azaldığı gözlemlenmektedir.






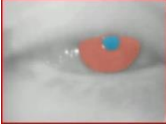

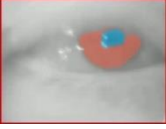



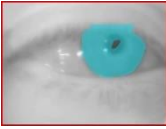
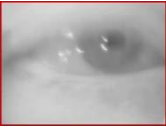



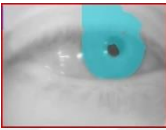





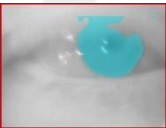


YOLO modellerinin performansları değerlendirildiğinde ise farklı sonuçlar gözlenmiştir. İlk karede, göz kapağının hafif kapalı olduğu durumda üç model de iris ve pupil bölgelerini başarılı bir şekilde tespit edebilmiştir. Ancak göz kapağının daha fazla kapandığı karelerde modeller genel olarak pupil bölgesini tespit edememiş, iris segmentasyonunda ise eksik ya da hatalı sonuçlar üretmiştir. Bazı karelerde modeller pupil bölgesini iris olarak yanlış sınıflandırmıştır.

	Kare 1	Kare 2	Kare 3	Kare 4	Kare 5
Orijinal Kare					
SAM2					
YOLOv8					
YOLOv11					
YOLOv12					

Şekil 5.6: Düşük göz kapağı ve sınırlı görüş senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.

YOLOv12-seg modelinde, diğer YOLO modellerinden farklı olarak, son karede göz kapağının oldukça kısık olduğu durumda bile iris bölgesinin kısmen tespit edildiği görülmektedir. Bu sonuçlar, YOLO modellerinin SAM2'ye göre düşük göz kapağı kaynaklı görünürlük kısıtlamalarından daha fazla etkilendiğini ortaya koymaktadır. Bu durum, SAM2'nin bu tür zorlu koşullarda daha dayanıklı olduğunu ve klinik uygulamalarda güvenilirliğini artırabileceğini göstermektedir.

3. Kısmen Kapalı Göz Kapağı ile Şiddetli Nistagmus: Bu senaryoda, hastanın göz kapağı kısmen kapalı olup, aynı zamanda güçlü bir nistagmus hareketi sergilemektedir. Göz kapağının kısık olması, segmentasyon modellerinin göz bebeğini net bir şekilde ayırt etmesini zorlaştırmaktadır. Geleneksel segmentasyon yöntemleri, göz kapağına ait pikselleri göz bebeği ile karıştırabilir ve hatalı tahminlerde bulunabilir.

	Kare 1	Kare 2	Kare 3	Kare 4	Kare 5
Orijinal Kare					
SAM2					
YOLOv8					
YOLOv11					
YOLOv12					

Şekil 5.7: Kısmen kapalı göz kapağı ve şiddetli nistagmus senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.

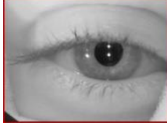



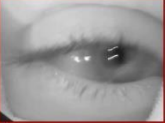











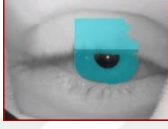








Görseller incelendiğinde, SAM2'nin çoğu karede göz bebeğini başarılı bir şekilde tespit edebildiği görülmektedir. İlk karelerde model, göz bebeğini belirgin şekilde maskeleyiş olup iris ve pupil sınırlarını doğru bir şekilde yakalamaktadır. Ancak, göz kapağı daha fazla kapandıkça segmentasyon maskesinin genişlediği ve pupili tespit edemediği fark edilmektedir. Bu durum, modelin dinamik sahnelerde belirli anlarda hatalar yapmasına rağmen genel olarak göz bebeğini takip edebildiğini göstermektedir.

YOLO modelleri ise iris ve pupil segmentasyonunda genel olarak tutarsız sonuçlar vermiştir. YOLOv8-seg ve YOLOv11-seg, ikinci karede nispeten doğru segmentasyon gerçekleştirmiş ancak diğer karelerde herhangi bir segmentasyon yapamamıştır. Ayrıca YOLOv11-seg, başarılı olduğu bu tek karede dahi YOLOv8-seg modeline göre daha düşük doğrulukta olup iris sınırlarında belirgin kaymalar göstermiştir.

YOLOv12-seg ise bir kare hariç tüm karelerde iris bölgesini geniş ve düzensiz olarak maskelemiş, pupil segmentasyonunu ise gerçekleştirememiştir. Bu sonuçlar, YOLO modellerinin kısmen kapalı göz kapağı ile şiddetli nistagmus içeren karmaşık sahnelerde zorluk yaşadığını ve SAM2'ye göre daha düşük performans gösterdiğini ortaya koymaktadır.

4. Yoğun Kirpik ve Kısıtlı Göz Açıklığı: Bu durumda, hastanın kirpikleri oldukça yoğun olup göz kapağı düşük konumdadır. Yoğun kirpikler, segmentasyon sürecinde ek engeller oluşturarak yanlış pozitif sonuçlara yol açabilmektedir. Geleneksel yöntemler, kirpiklerin oluşturduğu gölgeler nedeniyle göz bebeğinin sınırlarını hatalı belirleyebilir. Bu test senaryosunda, segmentasyon modelinin kirpiklerden etkilenmeden göz bebeğini ne kadar doğru tespit edebildiği incelenmiştir. Görseller incelendiğinde, SAM2 modelinin çoğu karede başarılı segmentasyon yaptığı görülmektedir. Ancak, son karede şiddetli nistagmus, yoğun kirpikler ve kısık göz kapağı gibi zorlu koşulların bir araya gelmesi modelin doğruluğunu etkilemiştir. Bu karede, model göz bebeğini segmentlemeye devam etse de sınırlarında kaymalar oluşmuş ve dış kısmını da segmentasyon maskesine dahil etmiştir. Buna rağmen, modelin kirpikleri veya göz kapağını segmentasyon alanına katmadığı gözlemlenmiştir. Bu durum, SAM2'nin belirli zorlu koşullarda dayanıklı olsa da çok karmaşık sahnelerde doğruluğunun azaldığını göstermektedir.

YOLOv8-seg sonuçları incelendiğinde, modelin ilk iki karede iris ve pupil segmentasyonunu kısmen doğru gerçekleştirdiği, ancak sonraki karelerde segmentasyon başarısının belirgin şekilde düştüğü gözlemlenmiştir. İlk karede hem iris hem de pupil doğru şekilde tespit edilmişken, ikinci karede yalnızca iris tespit edilebilmiştir.

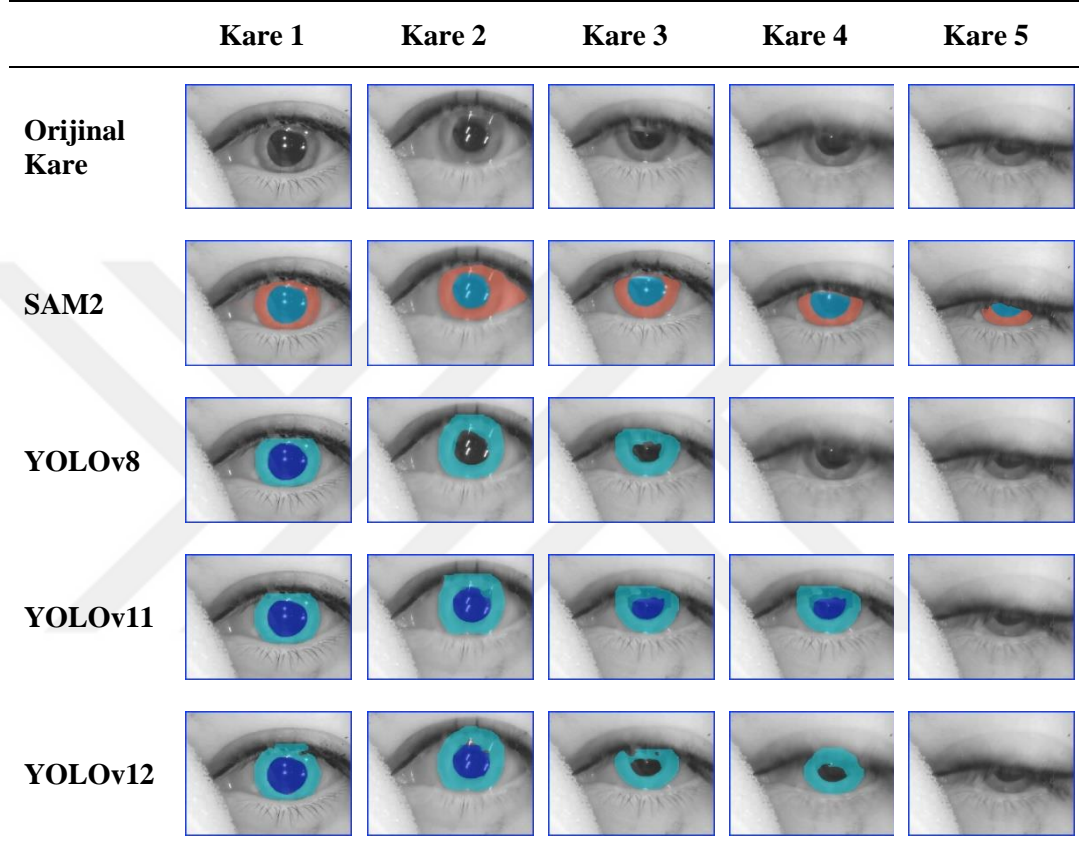
	Kare 1	Kare 2	Kare 3	Kare 4	Kare 5
Orijinal Kare					
SAM2					
YOLOv8					
YOLOv11					
YOLOv12					

Şekil 5.8: Yoğun kirpik ve sınırlı göz açıklığı senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.

Göz kapağının daha fazla kapandığı ve nistagmus hareketlerinin şiddetlendiği karelerde ise modelin segmentasyon yapamadığı görülmektedir. YOLOv11-seg ise yalnızca ilk karede başarılı bir segmentasyon gerçekleştirmiş, ancak sonraki karelerde segmentasyon hataları ve eksiklikler gözlemlenmiştir.

İkinci karede iris sınırlarında düzensizlikler oluşurken, sonraki karelerde modelin segmentasyon yapamadığı veya hatalı bölgelere maskeleye uyguladığı dikkat çekmektedir. YOLOv12-seg ise ilk iki karede iris ve pupil segmentasyonunu başarılı bir şekilde gerçekleştirmiştir. Ancak, göz kapağının daha kısık olduğu ve şiddetli nistagmus hareketlerinin bulunduğu karelerde segmentasyon doğruluğunun düştüğü gözlemlenmiştir. Son karelerde modelin iris sınırlarında belirgin kaymalar ve düzensizlikler olduğu, pupil segmentasyonunun ise tespit edilemediği görülmektedir.

5. Makyajlı Göz: Göz çevresine uygulanan makyaj, özellikle eyeliner veya koyu far gibi belirgin makyaj türleri, segmentasyon modelinin göz bebeği sınırlarını tespit etmesini zorlaştırabilir. Yoğun makyaj, göz bebeğinin kontrastını değiştirebilir ve modelin göz bebeğini yanlış tanımlamasına sebep olabilir. Bu senaryo, segmentasyon modelinin farklı ışık yansımaları ve kontrast değişiklikleri altında nasıl performans gösterdiğini değerlendirmek için seçilmiştir.



Şekil 5. 9: Makyaj kaynaklı zorluk senaryosu altında YOLOv8, YOLOv11, YOLOv12 ve SAM2 modellerinin segmentasyon sonuçlarının karşılaştırılması.

Görsel çıktılar incelendiğinde, SAM2 modelinin göz bebeğini başarılı bir şekilde tespit ettiği ve segmentasyon maskesini büyük ölçüde doğru konumlandığı görülmektedir. Özellikle koyu makyaj bulunan göz çevresinde, modelin kirpikleri segmentasyona dahil etmediği ve göz bebeğini kirpik aralarından doğru şekilde belirleyebildiği gözlemlenmiştir.

Model göz bebeğini net bir şekilde yakalamış ve maskeyi tutarlı bir biçimde uygulamıştır. Ancak, ikinci karede modelin segmentasyon maskesinin genişleyerek göz çevresindeki bazı bölgeleri de kapsadığı gözlemlenmiştir. Bu durum, aşırı gölge ve kontrast değişimlerinin model performansını sınırlayabileceğini göstermektedir.

YOLO modellerinin performansı incelendiğinde, iris ve pupil segmentasyonunda farklı seviyelerde doğruluk elde edildiği görülmektedir. YOLOv8-seg, iris ve pupil segmentasyonunda başlangıçta başarılı sonuçlar vermiş ancak göz kapağının daha fazla kapanmasıyla segmentasyon doğruluğunu kaybetmiştir. Özellikle pupil tespitinde eksiklikler meydana gelmiş, iris segmentasyonu ise göz kapağına doğru kaymalar göstermiştir.

Göz kapağının belirgin şekilde kısıldığı karelerde model segmentasyon yapamamıştır. YOLOv11-seg, iris segmentasyonunda genellikle tutarlı sonuçlar vermiş olsa da pupil segmentasyonunda başarısız kalmıştır. Model, hafif göz kapağı kapanmalarına karşı segmentasyon başarısını koruyabilmiş, ancak göz bebeğinin daha az görünür olduğu durumlarda doğruluğu azalmıştır.

Göz kapağının büyük ölçüde kapandığı durumda ise herhangi bir segmentasyon gerçekleşmemiştir. YOLOv12-seg, ilk karelerde iris segmentasyonunu gerçekleştirebilmiş ancak pupil tespitinde eksiklikler göstermiştir. Model, iris segmentasyonunu belirli oranlarda sürdürebilse de, göz kapağı ve kirpiklerin etkisi arttıkça segmentasyon doğruluğu azalmıştır. Göz kapağının aşırı kısıldığı durumda segmentasyon başarısız olmuştur.

Bu çalışmada, BPPV tanısında göz bebeği segmentasyonunun zorlu koşullarda nasıl performans gösterdiği incelenmiştir. SAM2 ve YOLO tabanlı modeller, farklı görsel engellerin (düşük göz kapağı, şiddetli nistagmus, yoğun kirpik, makyaj vb.) segmentasyon doğruluğu üzerindeki etkilerini anlamak amacıyla test edilmiştir. Deneysel sonuçlar, her iki yöntemin güçlü ve zayıf yönlerini ortaya koymuştur.

SAM2 modeli, tutarlı ve başarılı sonuçlar üretmiş, özellikle iris ve pupil segmentasyonunda yüksek doğruluk sağlamıştır. Nistagmus hareketlerinin şiddetlendiği ve göz kapağının kısıldığı durumlarda, model belirli hassasiyet kayıpları yaşamış ancak segmentasyon bütünlüğünü büyük ölçüde korumuştur. Yoğun kirpik veya koyu makyaj gibi kontrast değişiklikleri içeren senaryolarda, SAM2 genellikle göz bebeğini doğru bir şekilde tespit etmeyi başarmıştır. Ancak, aşırı düşük göz kapağı

gibi pupilin neredeyse tamamen kapanmasına neden olan ekstrem koşullarda, modelin segmentasyon doğruluğunda düşüş yaşadığı gözlenmiştir.

YOLO tabanlı modellerin performansı, iris segmentasyonunda genellikle istikrarlı olsa da, pupil segmentasyonunda değişken sonuçlar vermiştir. YOLOv8-seg, iris segmentasyonunu büyük ölçüde başarılı şekilde gerçekleştirse de pupil segmentasyonunda eksiklikler göstermiştir. YOLOv11-seg, iris tespiti konusunda kısmen başarılı olmasına rağmen pupil segmentasyonunu neredeyse hiç gerçekleştirememiştir. YOLOv12-seg, pupil tespitinde diğer YOLO modellerine göre daha başarılı sonuçlar üretse de, iris segmentasyonunda daha fazla kayma ve düzensizlik meydana gelmiştir. Şiddetli nistagmus, göz kapağı kısıklığı ve yoğun kirpik gibi faktörler karşısında, YOLO modellerinin hassasiyetinin azaldığı ve belirli durumlarda segmentasyonun tamamen başarısız olduğu gözlenmiştir.

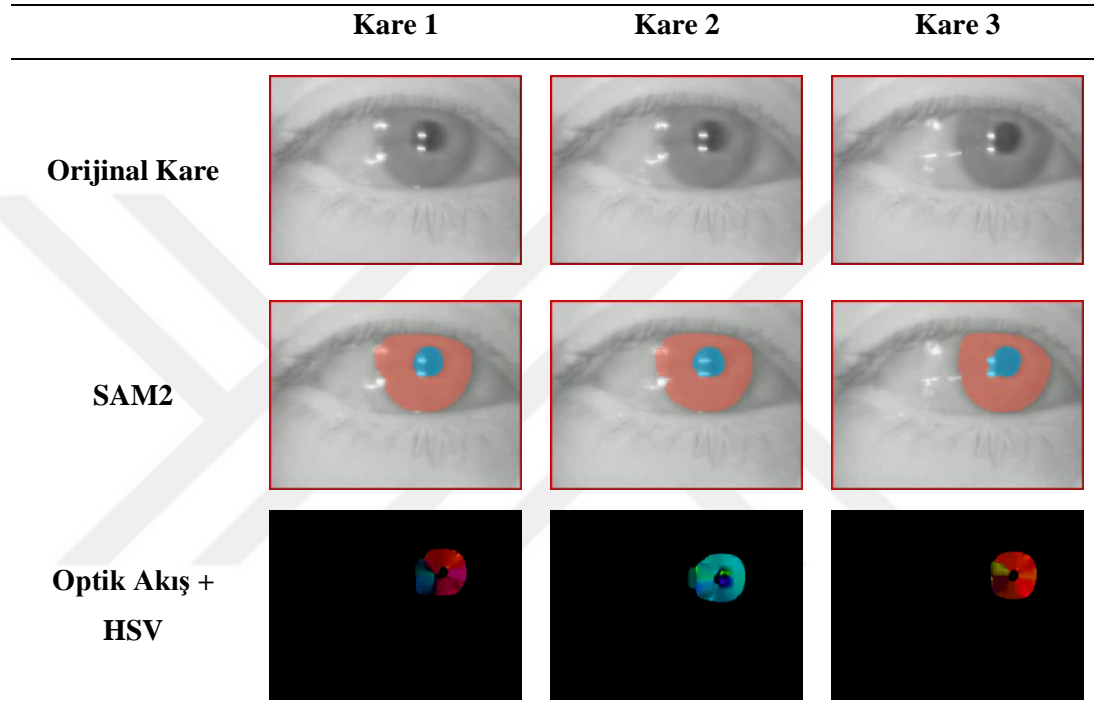
Genel olarak, SAM2'nin farklı zorluk seviyelerine karşı daha dirençli bir model olduğu, ancak ekstrem durumlarda hassasiyet kayıpları yaşadığı sonucuna varılmıştır. YOLO modelleri, bazı durumlarda iris tespitinde başarılı olsa da, pupil tespitinde daha düşük doğruluk oranlarına sahiptir. Bu durum, BPPV tanısına yönelik göz bebeği segmentasyonu için model seçiminde, kullanılan senaryonun ve ortam koşullarının dikkate alınması gerektiğini göstermektedir.

5.4. Optik Akış Analizi ve Hareket Örüntüleri

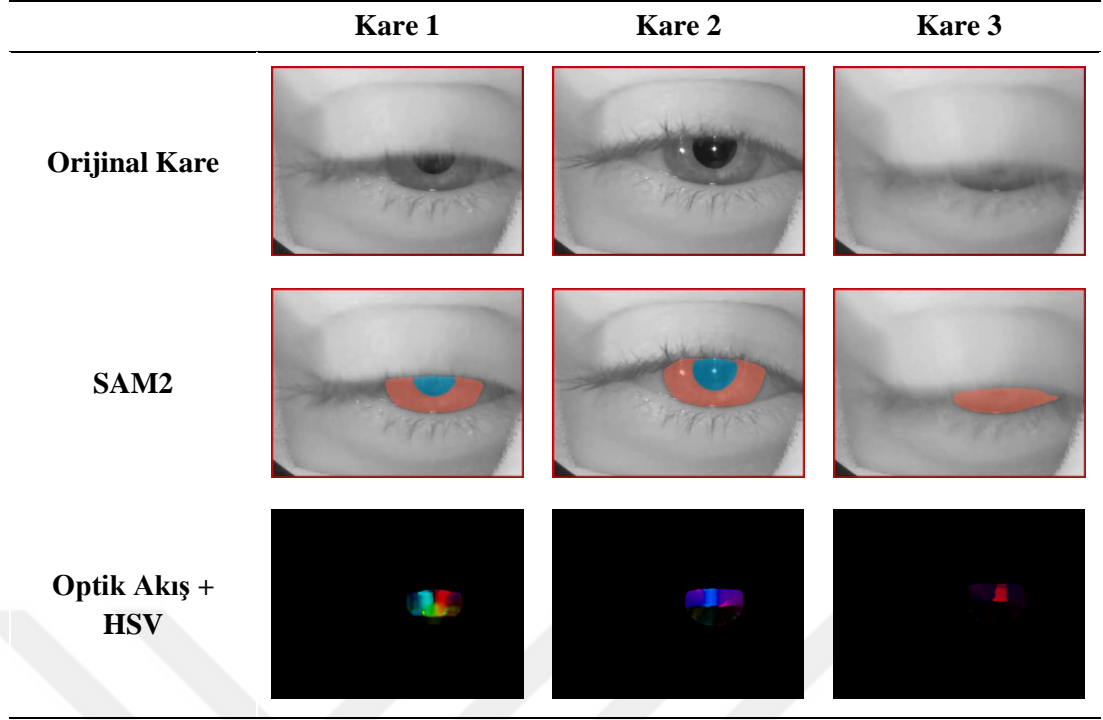
SAM2 segmentasyonunun doğruluğunu değerlendirdikten sonra, BPPV tanısında göz bebeği hareketlerinin önemini daha iyi analiz edebilmek amacıyla optik akış yöntemi uygulanmıştır. Bu analiz, yalnızca segmentasyon başarısını incelemek için değil, aynı zamanda zorlu koşullarda bile göz bebeğinin tespit edilerek hareket bilgilerinin değerlendirilebileceğini göstermek amacıyla gerçekleştirilmiştir.

Segmentasyon doğruluğu büyük önem taşımaktadır; çünkü göz bebeğinin en küçük hareketi bile tanı açısından kıymetli bilgiler sunabilir. Ancak, bazı karelerde segmentasyon başarısı düştüğünde bu kareleri tamamen göz ardı etmek, önemli tanısal bilgilerin kaybolmasına neden olabilir. Bu nedenle, optik akış yöntemi kullanılarak, zorlu koşullarda dahi göz bebeği hareketinin tespit edilip analiz edilebileceği gösterilmiştir.

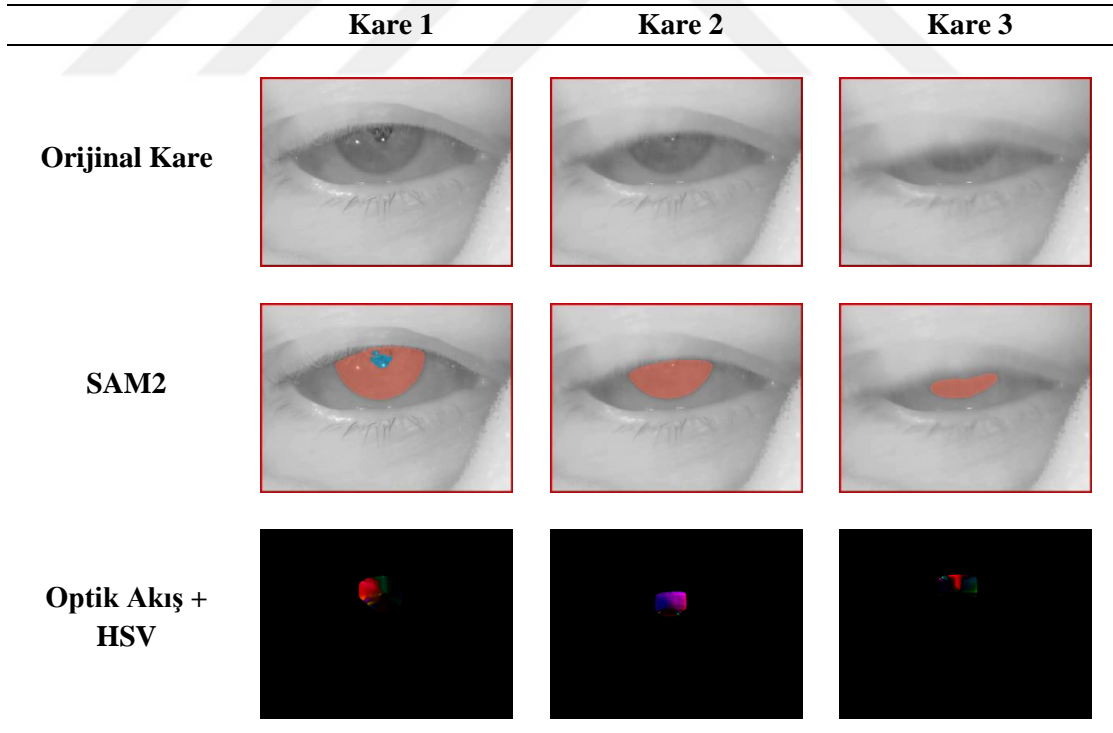
Bu süreçte YOLO modelleri yerine daha güvenilir segmentasyon sağlayan SAM2 çıktıları tercih edilmiştir. Çünkü zorlu senaryolarda göz bebeği tespiti başarısız olursa, hareket bilgisi de değerlendirilemez hale gelir. Bu analiz sürecinde, segmentasyon aşamasında YOLO modelleri yerine zorlu senaryolarda daha tutarlı sonuçlar üreten SAM2 modelinin çıktıları tercih edilmiştir. Çünkü göz bebeği doğru şekilde segmentlenmediğinde, ilgili karelerden elde edilecek hareket bilgileri de sağlıklı olmayacaktır.



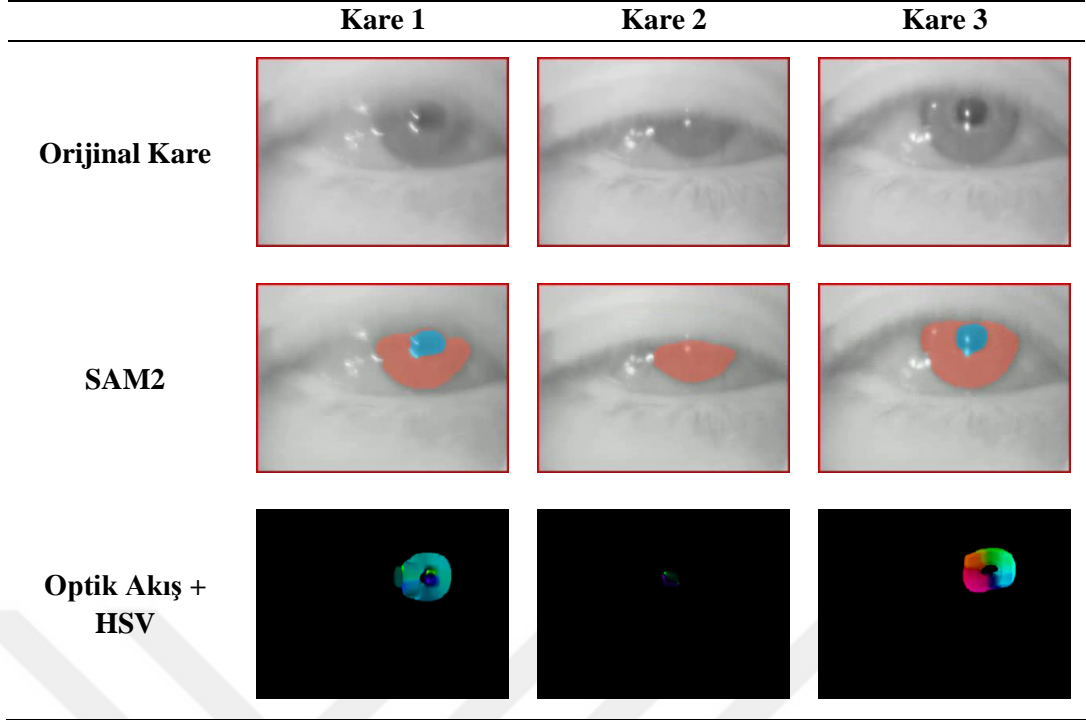
Şekil 5.10: Şiddetli nistagmus hareket senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.



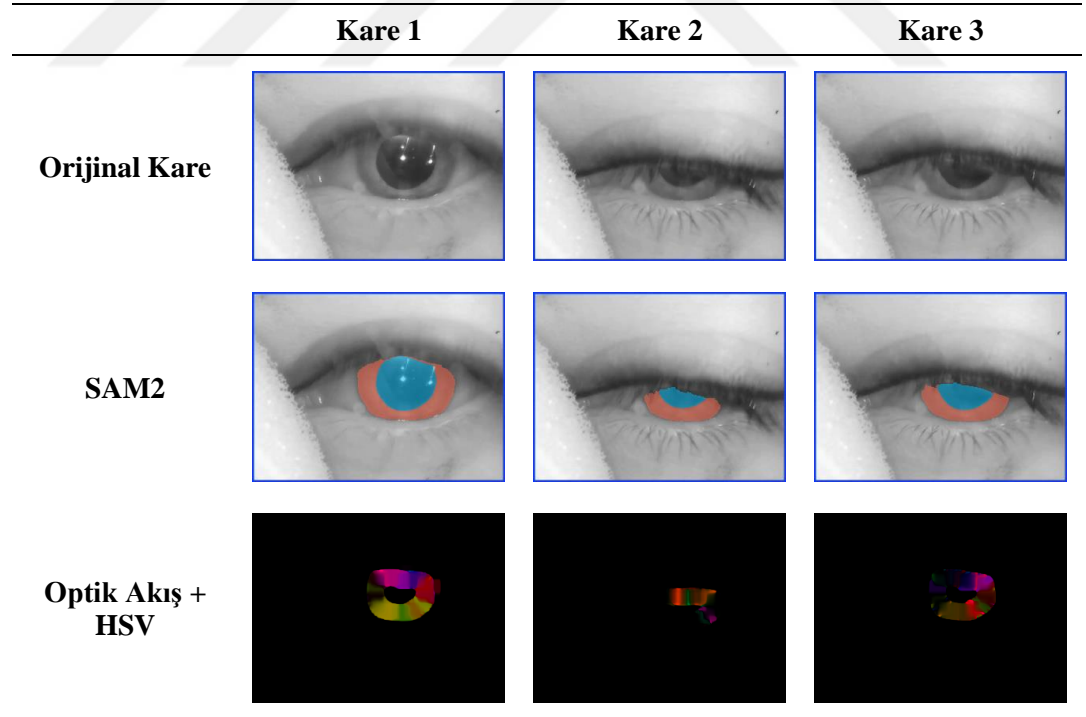
Şekil 5.11: Yoğun kirpik ve sınırlı göz açıklığı senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.



Şekil 5.12: Düşük göz kapağı ve sınırlı görüş senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.



Şekil 5.13: Kısmen kapalı göz kapağı ve şiddetli nistagmus senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.



Şekil 5.14: Makyaj kaynaklı zorluk senaryosu altında SAM2 maskeleri kullanılarak yapılan optik akış analiz sonuçları ve HSV renk kodlamasıyla görselleştirilmesi.

Segmentasyon sonrası, her videoda ardışık kareler arasındaki farklılıklara dayalı olarak Farneback yöntemiyle optik akış hesaplamaları gerçekleştirilmiştir. Böylece her pikselin yön ve hız bilgilerini içeren yoğun bir hareket haritası elde edilmiştir. Bu hareket vektörleri, HSV renk uzayına dönüştürülerek görselleştirilmiş; renk tonu hareketin yönünü, renk doygunluğu ve parlaklık ise şiddetini temsil edecek şekilde yorumlanabilir hale getirilmiştir. Tüm senaryolarda aynı işlem adımları izlenmiştir. Her bir senaryoya ait videolardan örnek kareler seçilerek, bu kareler üzerinden üçer farklı hareket görseli oluşturulmuştur.

Elde edilen sonuçlar, SAM2 ile segmentasyonun sağladığı avantaj sayesinde, nistagmus hareket yönünün düşük görünürlük koşullarında bile belirlenebileceğini göstermektedir. Bu, göz bebeğinin kısa süreliğine dahi görünür olduğu durumlarda, hareket yönünün doğru şekilde tespit edilmesini sağlayarak tanısal süreci destekleyebilir.

Önerilen yöntem sayesinde göz bebeği, en zorlu senaryolarda dahi başarılı bir şekilde segmentlenmiş ve optik akış analizi ile nistagmus hareketlerinin belirlenmesine olanak sağlanmıştır. Bu analiz, BPPV tanısında kritik olan göz bebeği hareketlerinin, zorlu koşullarda bile takip edilebildiğini ve tanıya katkı sağlayabileceğini ortaya koymaktadır.

5.5. CNN-GRU Sınıflandırma Performansı

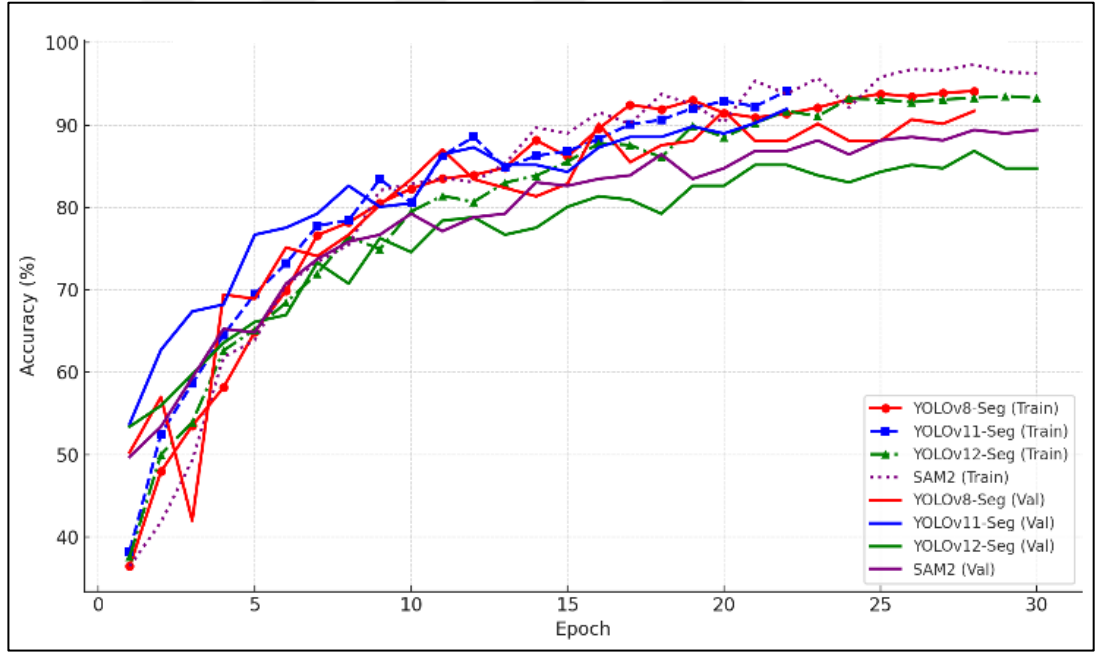
Segmentasyon adımının ardından, optik akış verileri kullanılarak dört farklı nistagmus yönü için sınıflandırma gerçekleştirilmiştir. Bu amaçla CNN-GRU mimarisi uygulanmış olup, modelin yapısı önceki bölümlerde ayrıntılı şekilde sunulmuştur. Model, 75 epoch boyunca eğitilmiş; early stopping mekanizması için patience=8 değeri belirlenmiştir. Optimizasyon için öğrenme oranı $1e-3$ olarak ayarlanmıştır.

Sınıflandırma performansı doğruluk, precision, recall ve F1-score gibi metrikler üzerinden değerlendirilmiştir. Elde edilen sonuçlar, kullanılan segmentasyon yöntemine göre gruplandırılarak Tablo 5.4'te sunulmuştur.

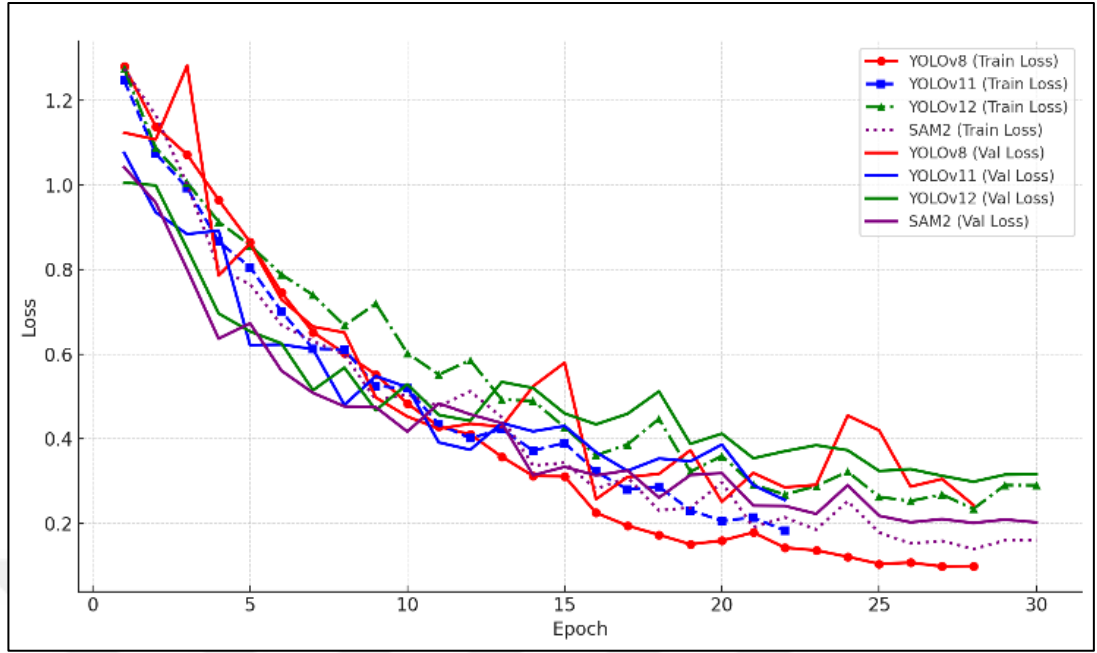
Tablo 5.4: Farklı segmentasyon yöntemleri kullanılarak elde edilen CNN-GRU modelinin sınıflandırma performansı.

Yöntem	Doğruluk	Precision	Recall	F1 Skoru
Yolov8-seg	0.85	0.85	0.85	0.85
Yolov11-seg	0.85	0.86	0.85	0.85
Yolov12-seg	0.89	0.89	0.89	0.89
SAM2	0.94	0.93	0.94	0.93

Sınıflandırma modellerinin başarısını değerlendirmek için yalnızca doğruluk (accuracy) metriğine odaklanmak, çoğu durumda yetersiz ve yanıltıcı olabilir. Özellikle sınıflar arasında dengesizlik olduğunda, doğruluk tek başına modelin genel performansını yansıtmakta eksik kalır. Bu nedenle; kesinlik (precision), duyarlılık (recall) ve F1 skoru (F1-score) gibi tamamlayıcı ölçütler de dikkate alınmalıdır.

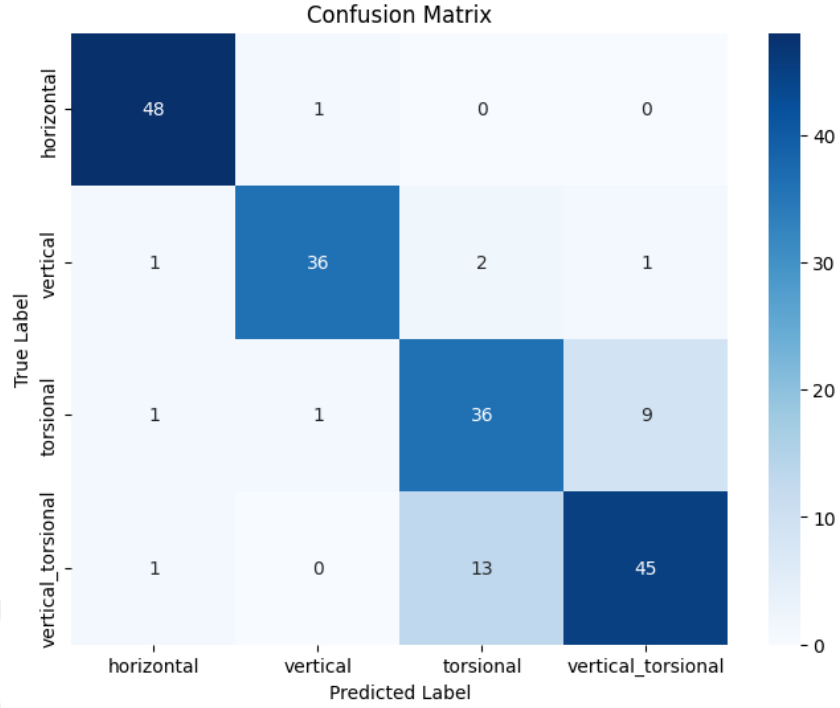


Şekil 5.15: Farklı segmentasyon yöntemleriyle (YOLOv8-Seg, YOLOv11-Seg, YOLOv12-Seg ve SAM2) eğitilen CNN-GRU modeline ait eğitim (train) ve doğrulama (validation-val) doğruluk eğrileri.

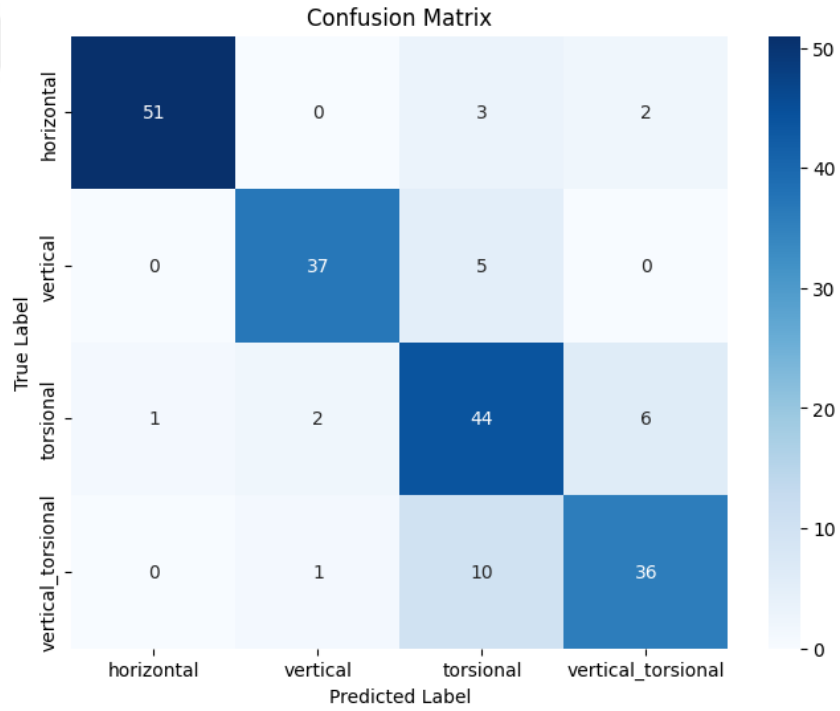


Şekil 5.16: Farklı segmentasyon yöntemleriyle (YOLOv8-Seg, YOLOv11-Seg, YOLOv12-Seg ve SAM2) eğitilen CNN-GRU modeline ait eğitim (train) ve doğrulama (validation-val) kayıp (loss) eğrileri.

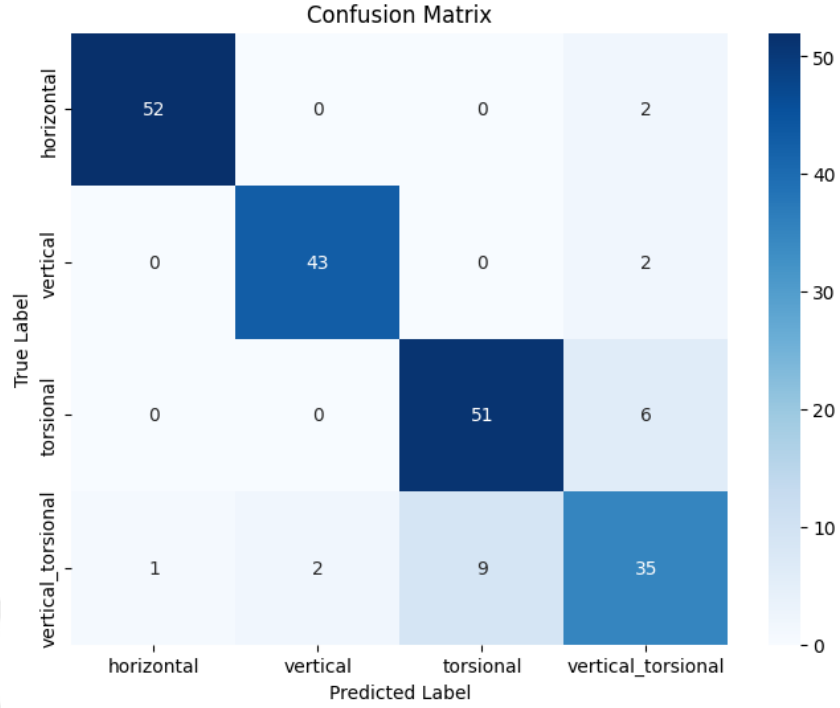
Elde edilen bulgular, segmentasyon yönteminin sınıflandırma başarımı üzerindeki etkisini ortaya koymaktadır. Dört farklı yöntemle yapılan karşılaştırmada, SAM2 tabanlı segmentasyon kullanıldığında modelin başarı ölçüt değerlerinin diğer yöntemlere kıyasla daha yüksek olduğu gözlemlenmiştir.



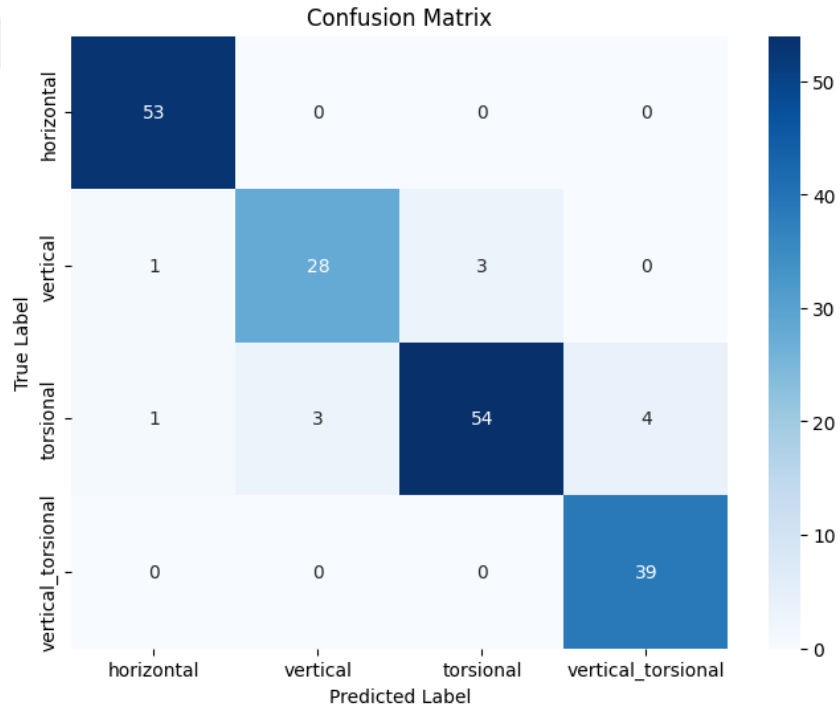
Şekil 5.17: YOLOv8-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.



Şekil 5.18: YOLOv11-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.



Şekil 5.19: YOLOv12-seg tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.



Şekil 5.20: SAM2 tabanlı segmentasyonla eğitilen CNN-GRU modeline ait karışıklık matrisi.

Bu genel başarı metriklerine ek olarak, sınıflandırma modelinin hangi sınıflarda ne ölçüde başarılı olduğunu daha ayrıntılı şekilde analiz edebilmek amacıyla her bir segmentasyon yöntemi için elde edilen confusion matrix (karışıklık matrisi) sonuçları sunulmuştur. Bu matrisler, modelin her bir sınıfı doğru ya da yanlış tahmin etme durumlarını açıkça ortaya koyarak, sınıf bazlı güçlü ve zayıf yönlerin belirlenmesine olanak tanımaktadır. Şekil 5.17–Şekil 5.20 arasında, dört farklı segmentasyon yöntemiyle eğitilen CNN-GRU modeline ait confusion matrix görselleştirmeleri karşılaştırmalı olarak verilmiştir.

Tablo 5.5: YOLOv8-seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.

	Precision	Recall	F1-skoru	Veri Sayısı
horizontal	0.94	0.98	0.96	49
vertical	0.95	0.90	0.92	40
torsional	0.71	0.77	0.73	47
vertical_torsional	0.82	0.76	0.79	59
Doğruluk			0.85	195
Makro ort.	0.85	0.85	0.85	195
Ağırlıklı ort.	0.85	0.85	0.85	195

Tablo 5.6: YOLOv11-Seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.

	Precision	Recall	F1-skoru	Veri Sayısı
horizontal	0.98	0.91	0.94	56
vertical	0.93	0.88	0.90	42
torsional	0.71	0.83	0.77	53
vertical_torsional	0.82	0.77	0.79	47
Doğruluk			0.85	198
Makro ort.	0.86	0.85	0.85	198
Ağırlıklı ort.	0.86	0.85	0.85	198

Tablo 5.7: YOLOv12-Seg segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.

	Precision	Recall	F1-skoru	Veri Sayısı
horizontal	0.98	0.96	0.97	54
vertical	0.96	0.96	0.96	45
torsional	0.85	0.89	0.87	57
vertical_torsional	0.78	0.74	0.76	47
Doğruluk			0.89	203
Makro ort.	0.86	0.85	0.89	203
Ağırlıklı ort.	0.86	0.85	0.89	203

Tablo 5.8: SAM2 segmentasyonu ile eğitilen CNN-GRU modeline ait sınıf bazlı metrikler.

	Precision	Recall	F1-skoru	Veri Sayısı
horizontal	0.96	1.00	0.98	53
vertical	0.90	0.88	0.89	32
torsional	0.95	0.87	0.91	62
vertical_torsional	0.91	1.00	0.95	39
Doğruluk			0.94	186
Makro ort.	0.86	0.85	0.93	186
Ağırlıklı ort.	0.86	0.85	0.93	186

CNN-GRU modelinin başarımını yalnızca genel metriklerle değil, her bir sınıf özelinde de değerlendirmek amacıyla sınıf bazlı precision, recall ve F1-score değerleri analiz edilmiştir. Bu analizler, özellikle nistagmus yönlerinin sınıflandırılmasında oluşabilecek dengesizliklerin tespiti ve modelin hangi sınıfta daha yüksek tutarlılıkla çalıştığının belirlenmesi açısından önemlidir.

Bu kapsamda, her bir segmentasyon yöntemi (YOLOv8-Seg, YOLOv11-Seg, YOLOv12-Seg ve SAM2) ile eğitilen CNN-GRU modellerine ait sınıf bazlı değerlendirme sonuçları Tablo 5.5–Tablo 5.8 arasında sunulmuştur. Bu tablolar

aracılığıyla, farklı segmentasyon yaklaşımlarının sınıf düzeyindeki etkileri karşılaştırmalı olarak incelenmiştir.

Karşılaştırılan çalışmalardaki sınıflarda belirtilen kısaltmalar şu anlama gelmektedir: SC, ASC ve PC sırasıyla superior, anterior ve posterior semisirküler kanalları ifade etmektedir. Geotropic ve apogeotropic terimleri, nistagmusun yerçekimi yönünde ya da tersine oluştuğunu belirtirken; CCTU ve CTU sırasıyla kanalliyazis ve kupulolitiyazis durumlarında başın yukarı çevrilmesini ifade eder. Sağ/sol ibareleri etkilenen tarafı gösterirken, “normal” patolojik bulgu olmayan durumu, “diğer” ise sınıflandırma dışında kalan özel durumları temsil etmektedir.

Tablo 5.9: BPPV'de nistagmus sınıflandırması için önerilen yöntemin derin öğrenme tabanlı yaklaşımlarla karşılaştırılması.

Referans#	Yöntem	Veri Sayısı	Sınıf Sayısı	Sınıflar	F1 Skoru
(Lim vd., 2019)	CNN Tabanlı Model	1005	8	SC Sağ/Sol Geotropic Sağ/Sol Apogeotropic Sağ/Sol ASC Sağ/Sol	0.79
(Pham vd., 2022)	Optik Akış + ResNext + BiGRU CNN + BiGRU	746	6	PC Sağ/Sol Geotropic Sağ/Sol Apogeotropic Sağ/Sol	0.90
(Wagle vd., 2022)	ResNet + VGG	435	2	Nystagmus Normal	0.82
(Kong vd., 2023)	Optik Akış + Iris Segmentasyonu ConvNext +LSTM	1060	4	CCTU CTU Sol Sağ	0.93
(Li ve Yang, 2023)	BiLSTM-GRU	21743	2	Vertical Diğer	0.92
Önerilen Çalışma (2025)	SAM2 + Optik Akış + CNN-GRU	1235	4	Horizontal Vertical Torsional Vertical + Torsional	0.93

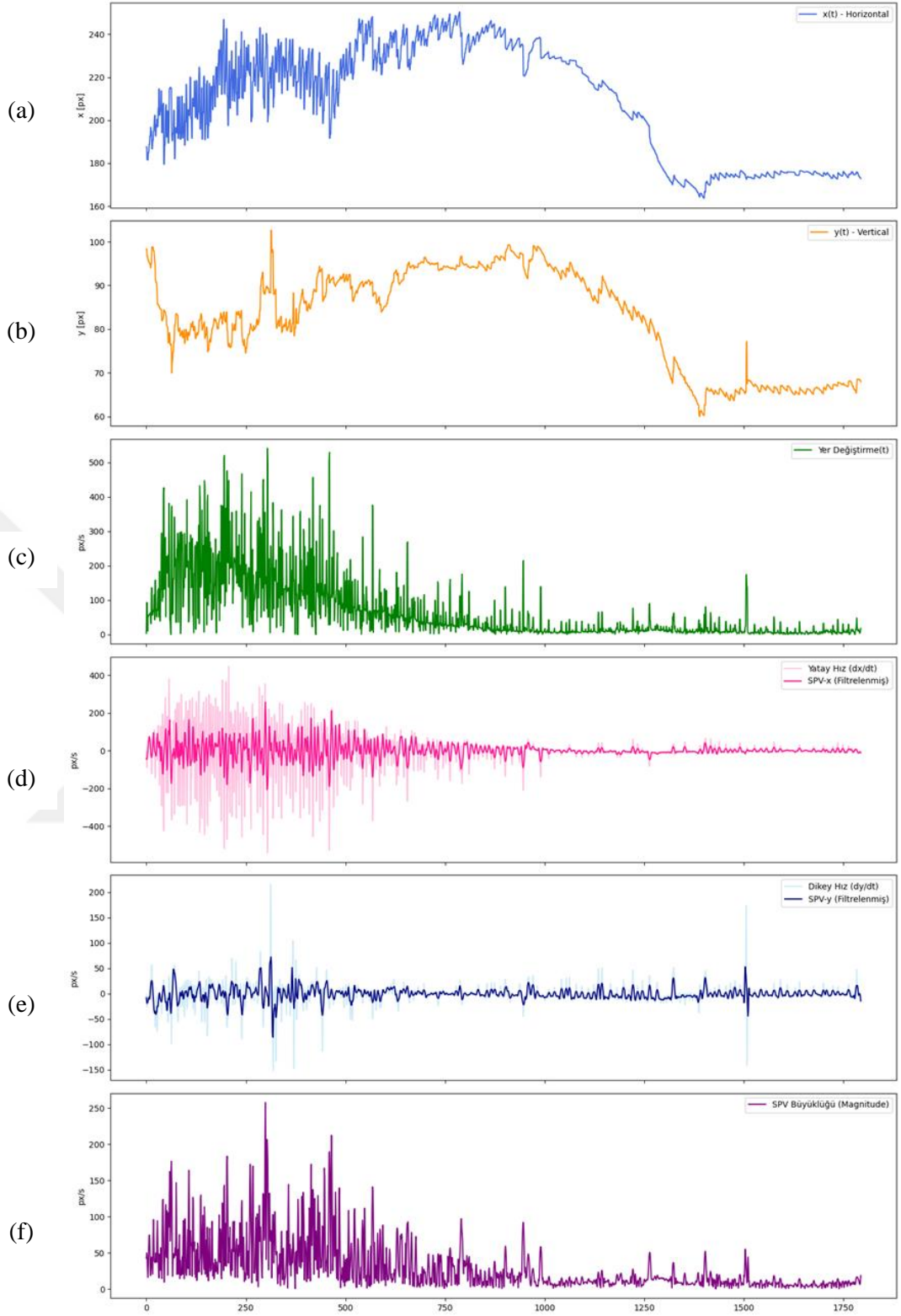
Tablo 5.9’da görüldüğü üzere, önerilen yöntem 0.93 F1 skoru elde ederek Kong vd. (2023) tarafından sunulan yöntemle benzer bir başarı göstermiştir. Bu tablo, aynı zamanda geliştirilen modelin dört farklı nistagmus türünü ayırt edebilme yeteneğini ortaya koymaktadır. Bu durum, yalnızca iki sınıf üzerinden değerlendirme yapan önceki çalışmalara kıyasla daha ayrıntılı ve klinik açıdan daha faydalı bir bakış açısı sunmaktadır. Veri kümesinin çok büyük olmamasına rağmen, güncel yapay zekâ tekniklerinin kullanılması, başarılı segmentasyon ve etkili sınıflandırma performansı yöntemin güçlü yönleri arasında yer almaktadır.

Öte yandan, her çalışmanın farklı sayıda ve türde nistagmus sınıfını hedef alması, F1 skorlarının doğrudan karşılaştırılmasını güçleştirebilir. Bu nedenle, yöntemler arasındaki farklılıkların göz önünde bulundurulması, sunulan başarı oranlarının daha sağlıklı yorumlanmasına katkı sağlayacaktır.

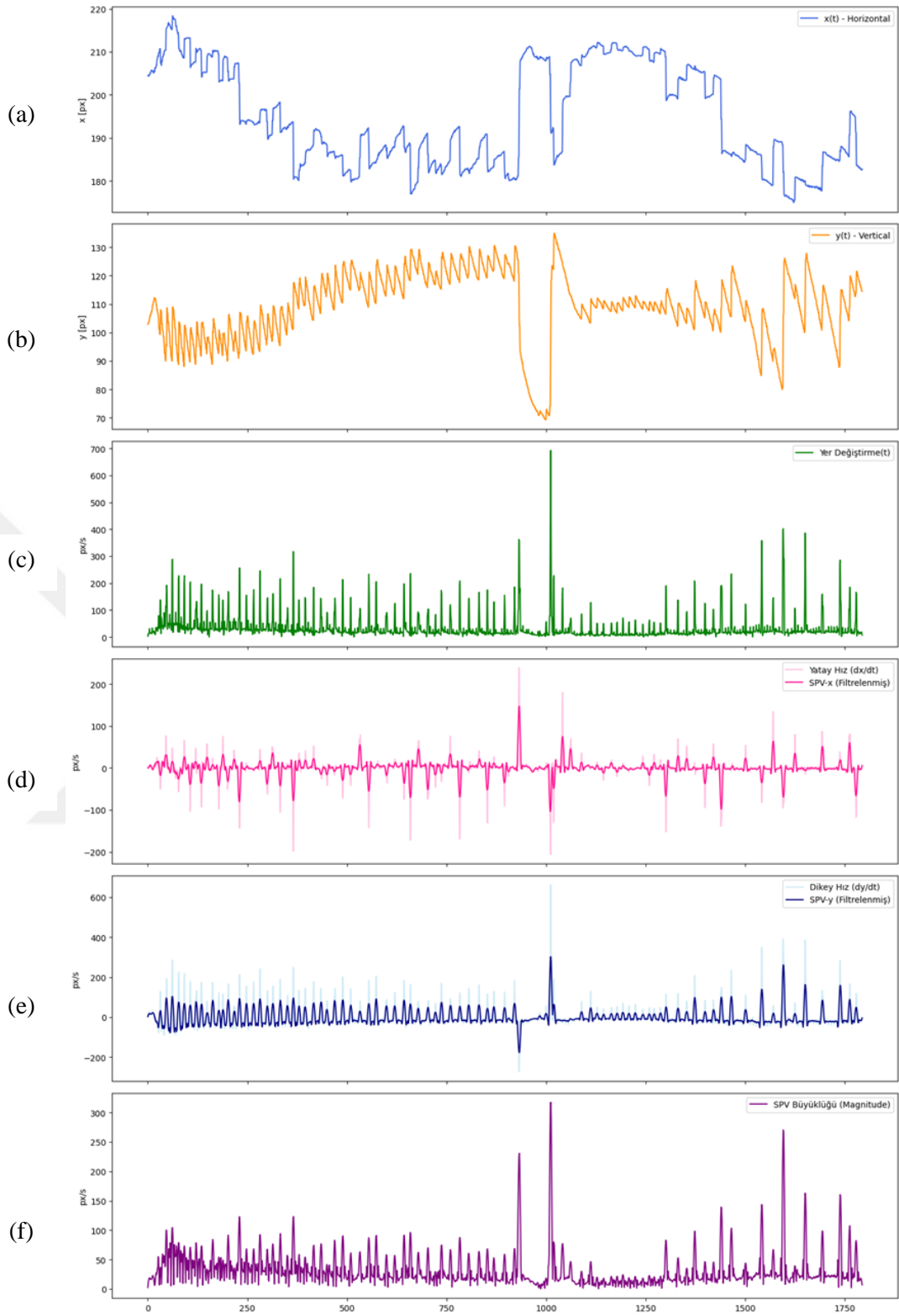
5.6. Bulanık Mantık Tabanlı Yorumlama ve Karar Desteği

Çalışmada kullanılan derin öğrenme modelleri yüksek sınıflandırma başarısı sunmasına karşın, “kara kutu” yapıları nedeniyle klinik uygulamalarda şeffaflık açısından sınırlılıklar içermektedir. Bu çalışmada kullanılan bulanık mantık yapısı, bu sınırlılığı aşmaya yönelik olarak önemli bir katkı sunmaktadır. Ortalama SPV değerlerine dayalı puanlama ve bulanık kural tabanlı karar yapısı, hem sistemin yorumlanabilirliğini artırmakta hem de hekimin sürece dâhil olmasını kolaylaştırmaktadır. Klinik uygulamalarda karar verme süreçleri yalnızca algoritmanın çıktısına değil, çıktının nasıl oluştuğuna da bağlıdır. Bu yönüyle bulanık mantık sistemi, hekimin güvenini artırmakta ve sistemin klinik pratiğe entegrasyonunu destekleyici bir rol üstlenmektedir.

Çalışmada öncelikle nistagmus hareketlerinin yatay (X eksen) ve dikey (Y eksen) bileşenleri ile yavaş faz hızı (SPV) değerleri incelenmiştir. Her bir eksen için elde edilen zaman serileri, grafikler aracılığıyla görselleştirilmiş; bu grafikler üzerinden hareketin yönü, şiddeti ve zaman içindeki değişimi değerlendirilmiştir. SPV grafikleri ise yavaş faz hızını daha net ortaya koyarak, klinik yorumlamaya ek değer katmaktadır. Analizler, segmentasyon adımı belirlenen göz bebeği merkez koordinatları temel alınarak gerçekleştirilmiş; elde edilen bulgular hem görsel hem de metinsel olarak raporlanmıştır.



Şekil 5.21: Horizontal nistagmus içeren video kaydına ait grafik analizleri.



Şekil 5.22: Vertical nistagmus içeren video kaydına ait grafik analizleri.

Şekil 5.21’de horizontal nistagmus, Şekil 5.22’de ise vertical nistagmus içeren video kayıtlarına ait analiz sonuçları sunulmaktadır. (a) alt grafiği, göz bebeği merkez koordinatlarının yatay bileşeninden elde edilen X eksen konum değişimini; (b) alt grafiği, dikey bileşenden elde edilen Y eksen konum değişimini; (c) alt grafiği, yatay ve dikey konum farklarının vektörel birleşiminden hesaplanan iki boyutlu yer değiştirme büyüklüğünü göstermektedir. (d) alt grafiğinde X eksenindeki ham hız (açık renkli eğri) ve filtrelenmiş yavaş faz hızı SPV-X (koyu renkli eğri) yer almakta; (e) alt grafiğinde Y eksenindeki ham hız (açık renkli eğri) ve filtrelenmiş yavaş faz hızı SPV-Y (koyu renkli eğri) gösterilmektedir. (f) alt grafiği ise yatay ve dikey SPV bileşenlerinin vektörel birleşiminden elde edilen SPV büyüklüğünü (magnitüde) ifade etmektedir.

X eksen grafikleri, göz bebeğinin yatay düzlemdeki konum değişimini zaman eksenine göre göstermektedir. Veriler, nistagmusun yatay bileşeninin manevra başlangıcında hızla ortaya çıktığını, belirli bir süre devam ettikten sonra ise kademeli olarak azaldığını göstermektedir. Yavaş fazlarda hareketin daha doğrusal bir eğilim izlediği, hızlı fazlarda ise keskin konum değişimlerinin meydana geldiği tespit edilmiştir. Bu görünüm, nistagmusun tipik yatay hareket örüntüsünü yansıtmakta ve SPV hesaplamalarının temelini oluşturmaktadır. Grafikten, düşük, orta ve yüksek hız bölgelerinin net biçimde ayırt edilebildiği görülmekte; bu da bulanık mantık modülünün doğru kategorilendirme yapabilmesine zemin hazırlamaktadır.

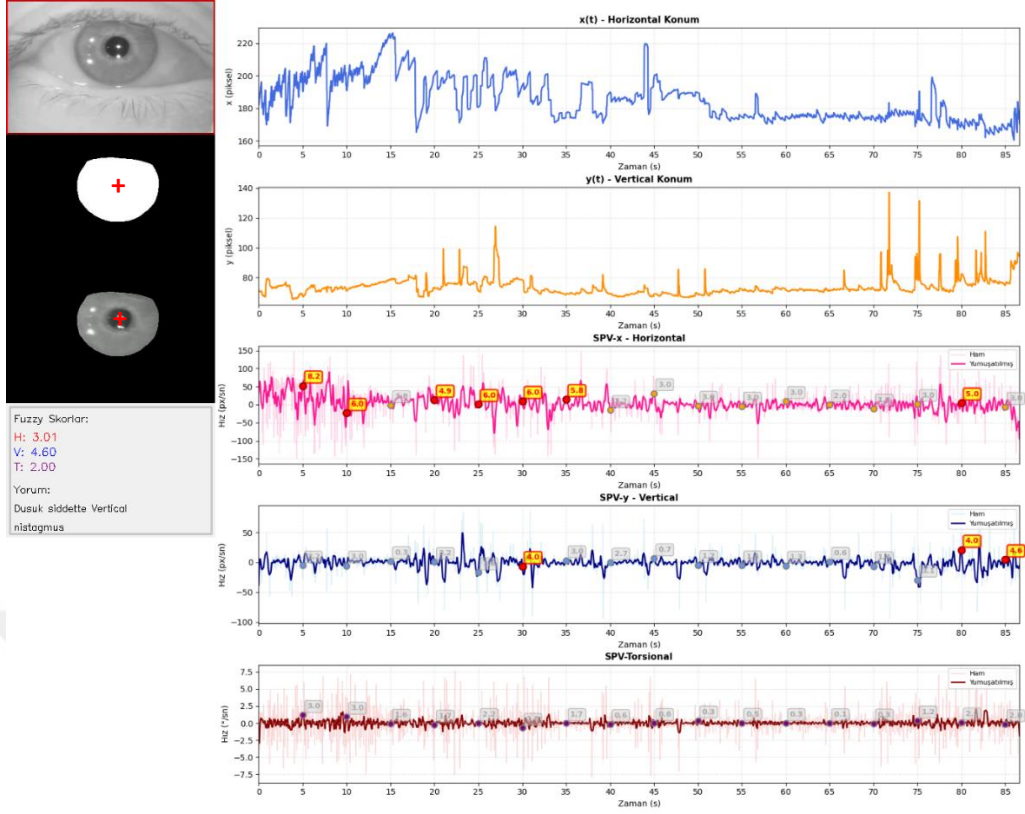
Y eksen grafikleri, göz bebeğinin dikey düzlemdeki konum değişimini zaman eksenine göre göstermektedir. Bulgular, dikey hareketin yatay harekete kıyasla daha sınırlı bir genlik sergilediğini; ancak belirli anlarda hızlanma ve yavaşlama dönemlerinin net şekilde ayırt edilebildiğini ortaya koymuştur. Başlangıçta daha yüksek konum değerlerinden başlayan hareketin, zamanla kademeli olarak azaldığı görülmektedir. Bu durum, manevra sonrasında dikey bileşenin yatay bileşene kıyasla daha erken sönmüldüğüne işaret etmektedir. Yavaş fazlarda hareketin daha yumuşak bir seyir izlediği, hızlı fazlarda ise kısa süreli ancak keskin değişimlerin meydana geldiği gözlenmiştir. Bu yapı, dikey nistagmusun karakteristik davranışını yansıtmakta ve SPV hesaplamalarının dikey eksen özelinde gerçekleştirilmesine olanak sağlamaktadır.

Konum deęişimlerinin yatay ve dikey bileşenleri birlikte deęerlendirildięinde, göz bebeęinin iki boyutlu yer deęiştirme büyüklüęü elde edilmektedir. Yer deęiştirme grafikleri, her iki eksenindeki konum farklarının vektörel olarak birleştirilmesiyle hesaplanan mesafenin zaman içindeki deęişimini göstermektedir. Bařlangıçta yüksek deęerlerin izlendięi bu grafik, ilerleyen zamanlarda mesafenin kademeli olarak azaldıęını ortaya koymaktadır. Yer deęiştirme büyüklüęü, nistagmusun genel hareket řiddetini tek bir ölçütle ifade etmesi açısından önem taşımakta; hem hızlı hem de yavaş fazların etkisini bütüncül bir şekilde yansıtmaktadır.

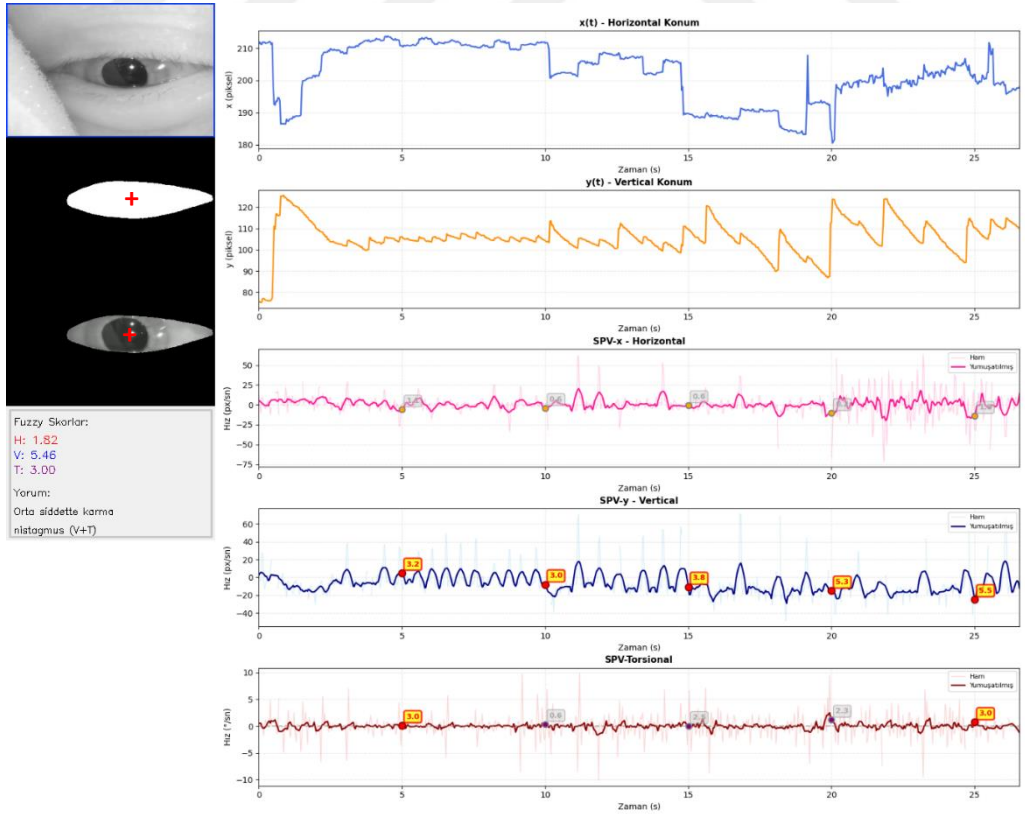
Hız analizinde, ilk olarak X eksenine ait yatay hız grafięi ele alınmaktadır. Açık renkli eęri, konum verilerinden türetilen ham hız deęerlerini; koyu renkli eęri ise filtreleme işleminde elde edilen yavaş faz hızını (SPV-X) göstermektedir. Yavaş faz hızının net biçimde ayrıştırılması, bulanık mantık tabanlı deęerlendirme süreci için temel giriş verisinin elde edilmesini saęlamaktadır. Zaman eksenini boyunca hız eęrisindeki deęişim, hızlı ve yavaş fazların dinamiklerini açıkça ortaya koymaktadır.

Dikey hız grafięi de benzer bir yapıya sahiptir. Açık renkli eęri, ham hız deęerlerini; koyu renkli eęri ise yavaş faz hızını (SPV-Y) temsil etmektedir. Dikey bileşende genliklerin yatay bileşene kıyasla daha düşük olduęu görülmekte, ancak zaman zaman belirgin hız artışları meydana gelmektedir. Yavaş fazların net şekilde ayrıştırılması, dikey bileşene ilişkin SPV hesaplamalarının da doęru biçimde yapılmasına imkân tanımaktadır. Son olarak, yatay ve dikey SPV bileşenlerinin vektörel olarak birleştirilmesiyle elde edilen SPV büyüklüęü (magnitude) grafikleri sunulmaktadır. Bu grafik, toplam yavaş faz hızını tek bir parametre altında ifade ederek nistagmusun genel řiddetini göstermektedir. Zaman içindeki deęişim, hem hareketin bařlangıcındaki hız karakterini hem de yavaşlayarak sona erme sürecini tek bir bakışta görmeyi saęlamaktadır.

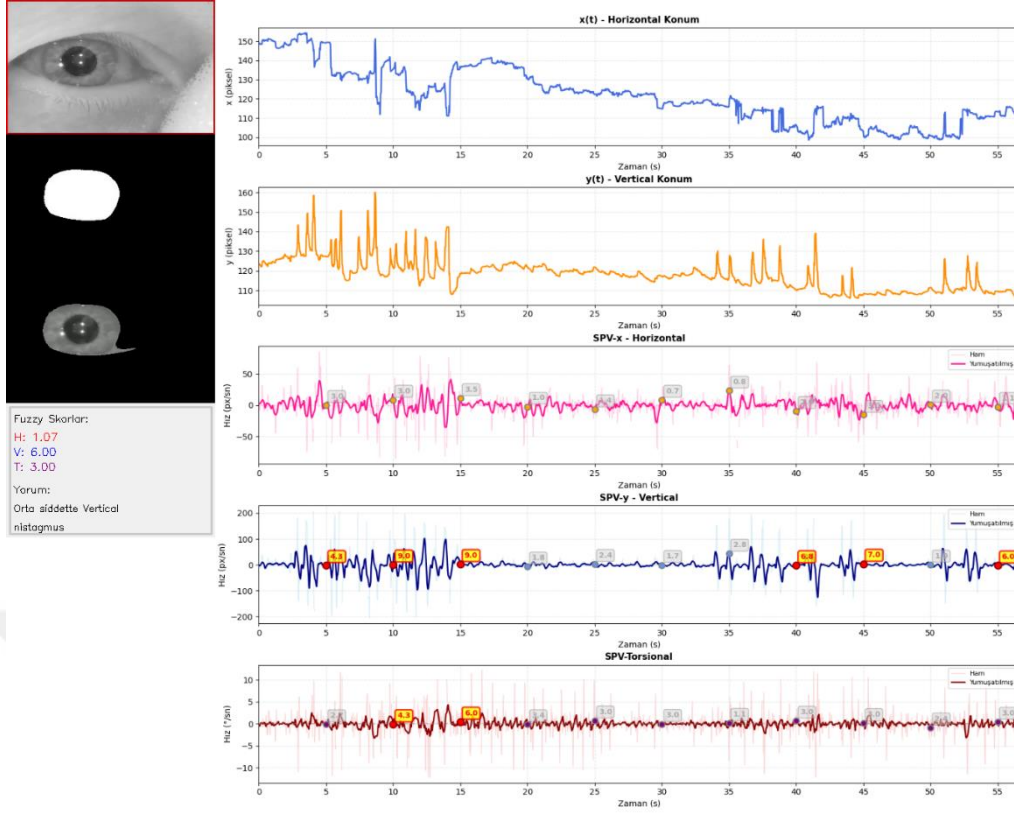
Sistemin nihai sürümünde torsiyonel bileşen de eklenmiş, böylece nistagmusun üç eksenindeki hareket karakteristikleri bulanık mantık yapısına entegre edilmiştir. Bu ekleme, özellikle karma nistagmus tiplerinin sınıflandırılmasında sistemin duyarlılıęını artırmıştır.



Şekil 5.23: Horizontal video analiz çıktısının son karesine ait ekran görüntüsü.



Şekil 5.24: Vertical video analiz çıktısının son karesine ait ekran görüntüsü.



Şekil 5.25: Torsiyonel video analiz çıktısının son karesine ait ekran görüntüsü.

Yapılan değerlendirmeler sonucunda, SPV değerlerine dayalı olarak geliştirilen bulanık mantık tabanlı karar sistemi, dört ana nistagmus yönü — horizontal, vertical, torsional ve vertical+torsional — için yorumlanabilir sınıflandırma sonuçları üretmiştir. Sistem, her 150 karede bir hesaplanan ortalama SPV değerini önceden tanımlanmış puan aralıklarına göre derecelendirmekte; bu puanları bulanık üyelik fonksiyonları aracılığıyla değerlendirerek kural tabanlı çıkarım mekanizmasıyla sınıf tahmini gerçekleştirmektedir. Kararların bu şekilde periyodik olarak üretilmesi, zaman içinde değişen nistagmus hareketlerinin dinamik olarak izlenmesine olanak sağlamakta; ayrıca her bir kararın ardındaki mantığın açık biçimde ortaya konması, sistemin klinik uygulamalarda güvenilirliğini ve yorumlanabilirliğini artırmaktadır. Derin öğrenme modellerinin yüksek doğruluklarına karşın sınırlı açıklanabilirliği dikkate alındığında, bu tür bulanık kural tabanlı yaklaşımların özellikle karmaşık ya da örneklem sayısı düşük klinik senaryolarda tamamlayıcı ve destekleyici bir rol üstlenebileceği değerlendirilmektedir.

Şekil 5.23–5.25’teki arayüz düzeninde, sol sütunda sırasıyla orijinal göz görüntüsü, SAM2 ile elde edilen maske, maske–orijinal görüntü bileşimi (iris dokusunu belirginleştiren görünüm) ve bulanık mantık modülünün ürettiği anlık skor ile sözel yorum sunulmaktadır. Sağ bölümde üstten alta doğru beş zaman serisi grafiği yer alır: $x(t)$ (yatay konum), $y(t)$ (dikey konum), SPV-X (yatay yavaş faz hızı), SPV-Y (dikey yavaş faz hızı) ve SPV-Torsiyonel (gözün kendi eksenini etrafındaki açısız yavaş faz hızı). SPV grafiklerinde pencere sonlarına karşılık gelen noktalayıcılar ve ilgili skor etiketleri görselleştirilerek karar anları açıkça işaretlenmiştir.

Sistem, kodda tanımlandığı üzere 5 saniyelik kayar pencereler kullanmaktadır (WINDOW = $5 \times \text{fps}$; örn. 30 fps için ≈ 150 kare). Her pencere sonunda hesaplanan SPV-X, SPV-Y ve SPV-T değerleri, tanımlı üyeli fonksiyonları ve kural tabanı üzerinden değerlendirilerek 0–10 aralığında sayısal skorlar ile bunlara karşılık gelen sözel klinik yorumlar üretilir. Bu periyodik ve dinamik güncelleme yapısı, niştagmusun zamana bağlı yön ve şiddet değişimlerinin bütüncül biçimde izlenmesine olanak tanımakta; yatay ve dikey bileşenlere eklenen torsiyonel bileşen sayesinde üç eksenli (horizontal–vertical–torsiyonel) yorumlanabilirlik sağlamaktadır. Böylece, hareketin yönü ve tipi yanında şiddeti de sayısal olarak ifade edilmekte, aynı zamanda kısa ve anlaşılır bir metin ile desteklenmektedir. Bu yapı, derin öğrenme modelinden elde edilen çıktıları anlamlandırarak klinik karar sürecine doğrudan katkı sağlayan açıklanabilir yapay zeka bileşeni olarak işlev görmektedir.

SONUÇ

Bu çalışmanın en önemli bulgularından biri, iris segmentasyonunun karmaşık klinik senaryolarda tanı sürecine sağladığı katkının net biçimde ortaya konmasıdır. Segmentasyonun ardından gelen sınıflandırma aşaması da nistagmus yönünün belirlenmesinde en az onun kadar kritik bir rol oynamıştır. Gerçek klinik koşullarda kısmen kapalı göz kapakları, yoğun kirpikler, ani göz hareketleri ve düşük kontrastlı görüntüler gibi faktörler segmentasyon doğruluğunu zorlaştırmaktadır. Bu nedenle modellerin yalnızca ideal koşullarda değil, klinik ortamlardaki değişken koşullar altında da değerlendirilmesi büyük önem taşımaktadır. Yapılan deneyler, SAM2 modelinin bu tür zorlayıcı durumlarda dahi tutarlı ve doğru segmentasyon maskeleri üretebildiğini göstermiştir.

Bu güçlü segmentasyon temelini üzerine inşa edilen optik akış analizi, iris hareketlerine ilişkin anlamlı zamansal veriler sağlayarak CNN-GRU tabanlı dört sınıflı nistagmus sınıflandırıcısına girdi oluşturmuştur. Bulgular, SAM2'nin yalnızca görsel kısıtlamalar altında yüksek doğrulukla çalışmakla kalmayıp, hareket analizi ve klinik yorumlama gibi aşağı akış görevlerinde de etkin biçimde kullanılabildiğini ortaya koymaktadır. YOLOv8-seg, YOLOv11-seg ve YOLOv12-seg modelleri de değerlendirilmiş; ancak özellikle göz kapağı örtüsü, hızlı nistagmus hareketleri ve düşük kontrastlı görüntülerde güvenilirlik sorunları yaşamışlardır. Bu nedenle, hareket analizinde süreklilik ve doğruluk sağlamak amacıyla yalnızca SAM2 çıktıları tercih edilmiştir.

Optik akış analizi, segmentasyonun sınırlı olduğu durumlarda dahi nistagmus takibinde anlamlı sonuçlar üretmiş; göz hareketlerinin yönü ve süresiyle ilgili verileri yakalayıp tanısal değeri artırmıştır. Segmentasyon kalitesinin, CNN-GRU sınıflayıcısının başarısını doğrudan etkilediği görülmüş; özellikle SAM2 maskeleri, optik akış hesaplamaları için daha kararlı bir zemin sağlayarak hareket desenlerinin kareler arasında güvenilir şekilde modellenmesine imkân tanımıştır.

Ayrıca, segmentasyon sonrası çıkarılan gözbebeği merkez hareketlerinden ve iris dokusundaki açısız değişimlerden elde edilen SPV değerleri hesaplanmış; bu değerler, yatay, dikey ve torsiyonel eksenler için tanımlanan kural tabanlı bulanık mantık sistemiyle birlikte yorumlanarak model çıktılarının açıklanabilirliği artırılmıştır. Böylece sistem, yalnızca doğru sınıflandırma yapmakla kalmamış, karar sürecini

görünür ve yorumlanabilir hâle getirerek klinik güveni de pekiştirmiştir. Elde edilen bulgular, segmentasyon ve sınıflandırma adımlarını bütünleşik biçimde optimize eden entegre karar destek sistemlerinin klinik tanı süreçlerinde yüksek değer taşıdığını göstermektedir. Bu yaklaşım, hareketin zaman içindeki dinamiklerini ve şiddetini etkili biçimde ortaya koyarak verilerin doğrudan klinik değerlendirmeye entegre edilmesini sağlamış; teknik doğruluk ile açıklanabilirliği bir araya getirmiştir.

Gelecek çalışmalarda, modelin genellenebilirliğinin daha geniş ve çeşitli bireyleri kapsayan klinik veri kümeleri üzerinde test edilmesi, farklı nistagmus tipleri için kural setlerinin genişletilmesi ve gerçek zamanlı performans optimizasyonlarıyla klinik iş akışlarına sorunsuz şekilde entegre edilmesi hedeflenmektedir.



KAYNAKÇA

- Afza, F., Khan, M. A., Sharif, M., Kadry, S., Manogaran, G., Saba, T., ... & Damaševičius, R. (2021). A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection. *Image and Vision Computing*, 106, 104090. <https://doi.org/10.1016/j.imavis.2020.104090>
- Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., ... & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information fusion*, 99, 101805. <https://doi.org/10.1016/j.inffus.2023.101805>
- Alonso-Fernandez, F., & Bigun, J. (2015). Near-infrared and visible-light periocular recognition with Gabor features using frequency-adaptive automatic eye detection. *IET Biometrics*, 4(2), 74-89. <https://doi.org/10.1049/iet-bmt.2014.0038>
- Anagnostou, E., Kouzi, I., & Spengos, K. (2015). Diagnosis and treatment of anterior-canal benign paroxysmal positional vertigo: a systematic review. *Journal of Clinical Neurology*, 11(3), 262-267. <https://doi.org/10.3988/jcn.2015.11.3.262>
- Athira, V., Geetha, P., Vinayakumar, R., & Soman, K. P. (2018). Deepairnet: Applying recurrent networks for air quality prediction. *Procedia computer science*, 132, 1394-1403. <https://doi.org/10.1016/j.procs.2018.05.068>
- Aw, S. T., Todd, M. J., Aw, G. E., McGarvie, L. A., & Halmagyi, G. M. (2005). Benign positional nystagmus: a study of its three-dimensional spatio-temporal characteristics. *Neurology*, 64(11), 1897-1905. <https://doi.org/10.1212/01.WNL.0000163545.57134.3D>
- Azimjonov, J., Özmen, A., & Varan, M. (2023). A vision-based real-time traffic flow monitoring system for road intersections. *Multimedia tools and applications*, 82(16), 25155-25174. <https://doi.org/10.1007/s11042-023-14418-w>

- Basha, S. S., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*, 378, 112-119. <https://doi.org/10.1016/j.neucom.2019.10.008>
- Bayram, O., Kucuk, H., Karabulut, B., & Canturk, M. (2025). The efficacy of automated repositioning chair in refractory benign paroxysmal positional vertigo-a pilot study. *Auris Nasus Larynx*, 52(2), 179-185. <https://doi.org/10.1016/j.anl.2025.01.012>
- Bhattacharyya, Neil, vd. "Clinical practice guideline: benign paroxysmal positional vertigo (update)." *Otolaryngology–Head and Neck Surgery* 156 (2017): S1-S47. <https://doi.org/10.1177/0194599816689667>
- Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6299-6308). <https://doi.org/10.48550/arXiv.1705.07750>
- Castillo, O., Valdez, F., Melin, P., & Ding, W. (2024). A survey on type-3 fuzzy logic systems and their control applications. *IEEE/CAA Journal of Automatica Sinica*, 11(8), 1744-1756. <https://doi.org/10.1109/JAS.2024.124530>
- Chandrakumar, T., & Kathirvel, R. J. I. J. E. R. T. (2016). Classifying diabetic retinopathy using deep learning architecture. *Int J Eng Res Technol*, 5(6), 19-24.
- Chen, T., Lu, A., Zhu, L., Ding, C., Yu, C., Ji, D., ... & Zang, Y. (2024). Sam2-adapter: Evaluating & adapting segment anything 2 in downstream tasks: Camouflage, shadow, medical image segmentation, and more. *arXiv preprint arXiv:2408.04579*. <https://doi.org/10.48550/arXiv.2408.04579>
- Chen, Y., Zhu, D., Shi, W., Zhang, G., Zhang, T., Zhang, X., & Li, J. (2023). Mfcflow: A motion feature compensated multi-frame recurrent network for optical flow estimation. In *Proceedings of the IEEE/CVF Winter conference on applications of computer vision* (pp. 5068-5077). <https://doi.org/10.1109/WACV56688.2023.00504>

- Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83, 187-205. <https://doi.org/10.1016/j.eswa.2017.04.030>
- Dogru, H. B., & Zengin, A. T. (2022). Fast-ICA Based Lane Detection Method for Autonomous Vehicles. In *2022 26th International Conference Electronics* (pp. 1-6). IEEE. <https://doi.org/10.1109/IEEECONF55059.2022.9810405>
- Dogru, H. B., Kaplan, A., Zengin, A. T., & Ozkul, T. (2024, April). Image-based Nystagmus Analysis in BPPV Disorders: Polar Coordinate System and Template Matching. In *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICMI60790.2024.10585886>
- Dong, W., Li, J., Yao, R., Li, C., Yuan, T., & Wang, L. (2016). Characterizing driving styles with deep learning. *arXiv preprint arXiv:1607.03611*. <https://doi.org/10.48550/arXiv.1607.03611>
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. <https://doi.org/10.48550/arXiv.1702.08608>
- Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. *Information systems frontiers*, 24(5), 1709-1734. <https://doi.org/10.1007/s10796-021-10186-w>
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639), 115-118. <https://doi.org/10.1038/nature21056>
- Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88, 303-338. <https://doi.org/10.1007/s11263-009-0275-4>
- Farneback, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Image Analysis: 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29–July 2, 2003 Proceedings 13* (pp. 363-370). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45103-X_50

- Fife, T. D., Iverson, D. J., Lempert, T., Furman, J. M., Baloh, R. W., Tusa, R. J., ... & Gronseth, G. S. (2008). Practice parameter: therapies for benign paroxysmal positional vertigo (an evidence-based review):[RETIRED] Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, 70(22), 2067-2074. <https://doi.org/10.1212/01.wnl.0000313378.77444.ac>
- Georgiev, P., Bhattacharya, S., Lane, N. D., & Mascolo, C. (2017). Low-resource multi-task audio sensing for mobile and embedded devices via shared deep neural network representations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3), 1-19. <https://doi.org/10.1145/3131895>
- Glenn Jocher and Jing Qiu. Ultralytics yolo11, 2024.
- Gonzales, R. C., & Wintz, P. (1987). *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc..
- Goswami, G., Bhardwaj, R., Singh, R., & Vatsa, M. (2014, September). MDLFace: Memorability augmented deep learning for video face recognition. In *IEEE international joint conference on biometrics* (pp. 1-7). IEEE. <https://doi.org/10.1109/BTAS.2014.6996299>
- Halmágyi, G. M., Akdal, G., Welgampola, M. S., & Wang, C. (2023). Neurological update: neuro-otology 2023. *Journal of Neurology*, 270(12), 6170-6192. <https://doi.org/10.1007/s00415-023-11922-9>
- Halmstad University Biometrics: Iris Segmentation Groundtruth Database. <https://github.com/HalmstadUniversityBiometrics/Iris-Segmentation-Groundtruth-Database>. Accessed on 20 April 2025 (2025)
- Hashemi, N. S., Aghdam, R. B., Ghiasi, A. S. B., & Fatemi, P. (2016). Template matching advances and applications in image analysis. arXiv preprint arXiv:1610.07231. <https://doi.org/10.48550/arXiv.1610.07231>
- Hizal, E., Jafarov, S., Erbek, S. H., & Ozluoglu, L. N. (2022). Clinical interpretation of positional nystagmus provoked by both Dix-Hallpike and supine head-roll tests. *The Journal of International Advanced Otolaryngology*, 18(4), 334. <https://doi.org/10.5152/iao.2022.21461>

- Hofbauer, H., Alonso-Fernandez, F., Wild, P., Bigun, J., & Uhl, A. (2014, August). A ground truth for iris segmentation. In *2014 22nd international conference on pattern recognition* (pp. 527-532). IEEE. <https://doi.org/10.1109/ICPR.2014.101>
- Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, *17*(1-3), 185-203. [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2)
- Institute of Automation, Chinese Academy of Sciences: CASIA-IrisV3. <http://www.cbsr.ia.ac.cn/IrisDatabase>. Accessed on 20 April 2025 (2025)
- Jain, A., Koppula, H. S., Raghavan, B., Soh, S., & Saxena, A. (2015). Car that knows before you do: Anticipating maneuvers via learning temporal driving models. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3182-3190).
- Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). Yolo evolution: A comprehensive benchmark and architectural review of yolov12, yolo11, and their previous versions. *Yolo11, and Their Previous Versions*. <http://doi.org/10.13140/RG.2.2.15952.83201>
- Jiaying, Z., & Hao, T. (2025). SAM2 for Image and Video Segmentation: A Comprehensive Survey. *arXiv preprint arXiv:2503.12781*. <https://doi.org/10.48550/arXiv.2503.12781>
- Jin, N., Mavromatis, S., Sequeira, J., & Curcio, S. (2020). A robust method of eye torsion measurement for medical applications. *Information*, *11*(9), 408. <https://doi.org/10.3390/info11090408>
- Kabade, V., Hooda, R., Raj, C., Awan, Z., Young, A. S., Welgampola, M. S., & Prasad, M. (2021). Machine learning techniques for differential diagnosis of vertigo and dizziness: a review. *Sensors*, *21*(22), 7565. <https://doi.org/10.3390/s21227565>
- Kao, W. T. K., Parnes, L. S., & Chole, R. A. (2017). Otoconia and otolithic membrane fragments within the posterior semicircular canal in benign paroxysmal positional vertigo. *The Laryngoscope*, *127*(3), 709-714. <https://doi.org/10.1002/lary.26115>

- Khanam, R., & Hussain, M. (2024). Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725*.
<https://arxiv.org/abs/2410.17725>
- Kim, J. S., & Zee, D. S. (2014). Benign paroxysmal positional vertigo. *New England Journal of Medicine*, 370(12), 1138-1147.
<https://doi.org/10.1056/NEJMcp1309481>
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4015-4026).
<https://arxiv.org/abs/2304.02643>
- Kong, S., Huang, Z., Deng, W., Zhan, Y., Lv, J., & Cui, Y. (2023). Nystagmus patterns classification framework based on deep learning and optical flow. *Computers in Biology and Medicine*, 153, 106473.
<https://doi.org/10.1016/j.combiomed.2022.106473>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25. <https://doi.org/10.1145/3065386>
- Kumar, A., Raj, R., & Mohammadzadeh, A. (2024). Recent advancements in type-3 fuzzy logic systems: a comprehensive review. *IEEE Transactions on Emerging Topics in Computational Intelligence*.
<https://doi.org/10.1109/TETCI.2024.3433514>
- Labib, R. P. M. D., Hadi, S., Widayaka, P. D., & Faradisa, I. S. (2022). Convolutional neural network for cataract maturity classification based LeNet. *Jurnal Varian*, 5(2), 97-106. <https://doi.org/10.30812/varian.v5i2.1629>
- Lechner, C., Taylor, R. L., Todd, C., MacDougall, H., Yavor, R., Halmagyi, G. M., & Welgampola, M. S. (2014). Causes and characteristics of horizontal positional nystagmus. *Journal of neurology*, 261, 1009-1017.
<https://doi.org/10.1007/s00415-013-7223-5>

- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551. <https://doi.org/10.1162/neco.1989.1.4.541>
- Li, H., & Yang, Z. (2023). Vertical Nystagmus Recognition Based on Deep Learning. *Sensors*, 23(3), 1592. <https://doi.org/10.3390/s23031592>
- Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., & Guo, Y. (2021). Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU). *Information Processing in Agriculture*, 8(1), 185-193. <https://doi.org/10.1016/j.inpa.2020.02.002>
- Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12), 6999-7019. <https://doi.org/10.1109/TNNLS.2021.3084827>
- Li, Z., Tang, W., Gao, S., Wang, Y., & Wang, S. (2024). Adapting SAM2 Model from Natural Images for Tooth Segmentation in Dental Panoramic X-Ray Images. *Entropy*, 26(12), 1059. <https://doi.org/10.3390/e26121059>
- Lim, E. C., Park, J. H., Jeon, H. J., Kim, H. J., Lee, H. J., Song, C. G., & Hong, S. K. (2019). Developing a diagnostic decision support system for benign paroxysmal positional vertigo using a deep-learning model. *Journal of clinical medicine*, 8(5), 633. <https://doi.org/10.3390/jcm8050633>
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer vision—ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13* (pp. 740-755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48
- Lu, H., Mao, Y., Li, J., & Zhu, L. (2024). Multimodal deep learning-based diagnostic model for BPPV. *BMC Medical Informatics and Decision Making*, 24(1), 82. <https://doi.org/10.1186/s12911-024-02438-x>

- Lu, W., Li, Z., Li, Y., Li, J., Chen, Z., Feng, Y., ... & Yin, S. (2022). A Deep Learning Model for Three-Dimensional Nystagmus Detection and Its Preliminary Application. *Frontiers in Neuroscience*, 16, 930028. <https://doi.org/10.3389/fnins.2022.930028>
- Lucas, B. D., & Kanade, T. (1981, August). An iterative image registration technique with an application to stereo vision. In *IJCAI'81: 7th international joint conference on Artificial intelligence* (Vol. 2, pp. 674-679).
- Lundén, J., & Koivunen, V. (2016, May). Deep learning for HRRP-based target recognition in multistatic radar systems. In *2016 IEEE Radar Conference (RadarConf)* (pp. 1-6). IEEE. <https://doi.org/10.1109/RADAR.2016.7485271>
- Ma, J., Kim, S., Li, F., Baharoon, M., Asakereh, R., Lyu, H., & Wang, B. (2024). Segment anything in medical images and videos: Benchmark and deployment. *arXiv preprint arXiv:2408.03322*. <https://doi.org/10.48550/arXiv.2408.03322>
- Meta AI Research: Segment Anything Model (SAM). <https://segment-anything.com>. Accessed on 20 April 2025 (2025). <https://doi.org/10.3390/electronics14091850>
- Milman, O., Yellin, D., & Aperstein, Y. (2025). Adapting SAM for Visible-Light Pupil Segmentation Baseline. *Electronics*, 14(9), 1850. <https://doi.org/10.3390/electronics14091850>
- Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 215232. <https://doi.org/10.3389/fpls.2016.01419>
- Newman-Toker, D. E., Tehrani, A. S. S., Mantokoudis, G., Pula, J. H., Guede, C. I., Kerber, K. A., ... & Kattah, J. C. (2013). Quantitative video-oculography to help diagnose stroke in acute vertigo and dizziness. <https://doi.org/10.1161/strokeaha.111.000033>
- Nie, L., Wang, M., Zhang, L., Yan, S., Zhang, B., & Chua, T. S. (2015). Disease inference from health-related questions via sparse deep learning. *IEEE Transactions on knowledge and Data Engineering*, 27(8), 2107-2119. <https://doi.org/10.1109/TKDE.2015.2399298>

- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25, pp. 15-24). San Francisco, CA, USA: Determination press.
- Ong, J. K., & Haslwanter, T. (2010). Measuring torsional eye movements by tracking stable iris features. *Journal of neuroscience methods*, 192(2), 261-267.
- OpenCV. *Optical Flow*. OpenCV Documentation. https://docs.opencv.org/3.4/d4/dee/tutorial_optical_flow.html (Eriřim Tarihi: 20 Nisan 2025)
- Otero-Millan, J., Roberts, D. C., Lasker, A., Zee, D. S., & Kheradmand, A. (2015). Knowing what the brain is seeing in three dimensions: A novel, noninvasive, sensitive, accurate, and low-noise technique for measuring ocular torsion. *Journal of vision*, 15(14), 11-11. <https://doi.org/10.1167/15.14.11>
- Özcan, K. (t.y.). *Convolutional Neural Networks (CNN) Nedir?* Medium. https://medium.com/@kubra.ozcan_98680/convolutional-neural-networks-cnn-nedir-a3a9d5318835 (Eriřim Tarihi: 20 Mayıs 2025)
- Ozkul vd., “AI Supported Robotic Diagnostic Treatment System for Vertigo”, <https://contest.techbriefs.com/2022/entries/medical/11956>, (25 Şubat 2023).
- Öztürk, B., Güleç, M., Deveci, T. N., & Güler, M. T. (2019). Benign paroksizmal pozisyonel vertigo: patofizyoloji, deęerlendirme ve tanılama. *Türk Odyoloji ve İřitme Arařtırmaları Dergisi*, 2(1), 18-28. <https://doi.org/10.34034/TJAHR.23170>
- Pan, E., Mei, X., Wang, Q., Ma, Y., & Ma, J. (2020). Spectral-spatial classification for hyperspectral image based on a single GRU. *Neurocomputing*, 387, 150-160. <https://doi.org/10.1016/j.neucom.2020.01.029>
- Pham, T. T., & Shen, Y. (2017). A deep causal inference approach to measuring the effects of forming group loans in online non-profit microfinance platform. *arXiv preprint arXiv:1706.02795*. <https://doi.org/10.48550/arXiv.1706.02795>

- Pham, T. X., Choi, J. W., Mina, R. J. L., Nguyen, T. X., Madjid, S. R., & Yoo, C. D. (2022). Lad: A hybrid deep learning system for benign paroxysmal positional vertigo disorders diagnostic. *IEEE Access*, *10*, 113995-114007. <https://doi.org/10.1109/ACCESS.2022.3215625>
- Piccialli, F., Di Somma, V., Giampaolo, F., Cuomo, S., & Fortino, G. (2021). A survey on deep learning in medicine: Why, how and when?. *Information Fusion*, *66*, 111-137. <https://doi.org/10.1016/j.inffus.2020.09.006>
- Pişkin, M. *OpenCV ile Renk Tespiti ve Obje Takibi*. Mesut Pişkin Blogu. <https://mesutpiskin.com/blog/opencv-renk-tespiti-obje-takibi.html> (Erişim Tarihi: 20 Nisan 2025)
- Poplin, R., Varadarajan, A. V., Blumer, K., Liu, Y., McConnell, M. V., Corrado, G. S., ... & Webster, D. R. (2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature biomedical engineering*, *2*(3), 158-164. <https://doi.org/10.1038/s41551-018-0195-0>
- Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, *14*(1), 6086. <http://dx.doi.org/10.1038/s41598-024-56706-x>
- Rashed, H., El Sallab, A., Yogamani, S., & ElHelw, M. (2019). Motion and depth augmented semantic segmentation for autonomous navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0). <https://doi.org/10.1109/CVPRW.2019.00049>
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2016). Deep learning for health informatics. *IEEE journal of biomedical and health informatics*, *21*(1), 4-21. <https://doi.org/10.1109/CVPRW.2019.00049>
- Ravi, N., Gabeur, V., Hu, Y. T., Hu, R., Ryali, C., Ma, T., ... & Feichtenhofer, C. (2024). Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*. <https://arxiv.org/abs/2408.00714>

- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788). <https://doi.org/10.1109/CVPR.2016.91>
- Ro, D. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15, 204-208. <https://doi.org/10.1145/361237.361242>
- Saatchi, R. (2024). Fuzzy logic concepts, developments and implementation. *Information*, 15(10). <https://doi.org/10.3390/info15100656>
- Sankar, M., Batri, K., & Parvathi, R. (2016). Earliest diabetic retinopathy classification using deep convolution neural networks. pdf. *Int. J. Adv. Eng. Technol*, 10, M9. <http://dx.doi.org/10.6084/M9.FIGSHARE.3407482.V1>
- Sapkota, R., & Karkee, M. (2024). Comparing YOLOv11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment. *arXiv preprint arXiv:2410.19869*. <https://doi.org/10.48550/arXiv.2410.19869>
- Sarda, A., Dixit, S., & Bhan, A. (2021, February). Object detection for autonomous driving using yolo [you only look once] algorithm. In *2021 Third international conference on intelligent communication technologies and virtual mobile networks (ICICV)* (pp. 1370-1374). IEEE <https://doi.org/10.1109/ICICV50876.2021.9388577>
- Sathyanarayanan, S., & Tantri, B. R. (2024). Confusion matrix-based performance evaluation metrics. *African Journal of Biomedical Research*, 4023-4031. <http://dx.doi.org/10.53555/AJBR.v27i4S.4345>
- Saxe, A., Nelli, S., & Summerfield, C. (2021). If deep learning is the answer, what is the question?. *Nature Reviews Neuroscience*, 22(1), 55-67. <https://doi.org/10.1038/s41583-020-00395-8>
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>

- Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*.
<https://doi.org/10.48550/arXiv.1312.6229>
- Shamsolmoali, P., Li, X., & Wang, R. (2019). Single image resolution enhancement by efficient dilated densely connected residual network. *Signal Processing: Image Communication*, 79, 13-23.
<https://doi.org/10.1016/j.image.2019.08.008>
- Shoman, M., Lanzaro, G., Sayed, T., & Gargoum, S. (2024). Autonomous Vehicle–Pedestrian Interaction Modeling Platform: A Case Study in Four Major Cities. *Journal of Transportation Engineering, Part A: Systems*, 150(9), 04024045.
<https://doi.org/10.1061/JTEPBS.TEENG-8097>
- Slama, A. B., Mouelhi, A., Sahli, H., Manoubi, S., Lahiani, R., Salah, M. B., ... & Sayadi, M. (2018). A new neural network method for peripheral vestibular disorder recognition using VNG parameter optimisation. *International Journal of Biomedical Engineering and Technology*, 27(4), 321-336.
<http://dx.doi.org/10.1504/IJBET.2018.10015346>
- Smith, A. R. (1978). Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3), 12-19. <https://doi.org/10.1145/965139.807361>
- Sohan, M., Sai Ram, T., & Rami Reddy, C. V. (2024). A review on yolov8 and its advancements. In *International Conference on Data Intelligence and Cognitive Informatics* (pp. 529-545). Springer, Singapore. http://dx.doi.org/10.1007/978-981-99-7962-2_39
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
<https://doi.org/10.48550/arXiv.1409.4842>
- Talmud, J. D., Coffey, R., Hsu, N. M., & Edemekong, P. F. (2017). Dix-Hallpike Maneuver.

- Tang, H., & Li, W. (2017). Advances in the diagnosis and treatment of benign paroxysmal positional vertigo. *Experimental and therapeutic medicine*, 14(3), 2424-2430. <https://doi.org/10.3892/etm.2017.4837>
- Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine learning and knowledge extraction*, 5(4), 1680-1716. <https://doi.org/10.3390/make5040083>
- Tian, Y., Ye, Q., & Doermann, D. (2025). Yolov12: Attention-centric real-time object detectors. *arXiv preprint arXiv:2502.12524*. <https://doi.org/10.48550/arXiv.2502.12524>
- Van Dam, V. S., Maas, B. D., Schermer, T. R., Van Benthem, P. P. G., & Brintjes, T. D. (2021). Two symptoms strongly suggest benign paroxysmal positional vertigo in a dizzy patient. *Frontiers in Neurology*, 11, 625776. <https://doi.org/10.3389/fneur.2020.625776>
- Vannucchi, P., Pecci, R., Giannoni, B., Di Giustino, F., Santimone, R., & Mengucci, A. (2015). Apogeotropic posterior semicircular canal benign paroxysmal positional vertigo: some clinical and therapeutic considerations. *Audiology Research*, 5(1), 130. <https://doi.org/10.4081/audiores.2015.130>
- Videnovic, J., Lukezic, A., & Kristan, M. (2025). A distractor-aware memory for visual object tracking with sam2. In *Proceedings of the Computer Vision and Pattern Recognition Conference* (pp. 24255-24264). <https://doi.org/10.48550/arXiv.2411.17576>
- Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76, 89-106. <https://doi.org/10.1016/j.inffus.2021.05.009>
- Wagle, N., Morkos, J., Liu, J., Reith, H., Greenstein, J., Gong, K., ... & Green, K. E. (2022). aEYE: a deep learning system for video nystagmus detection. *Frontiers in neurology*, 13, 963968. <https://doi.org/10.3389/fneur.2022.963968>
- Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, 429, 215-244. <https://doi.org/10.1016/j.neucom.2020.10.081>

- Xiong, X., Wu, Z., Tan, S., Li, W., Tang, F., Chen, Y., ... & Li, G. (2024). Sam2-unet: Segment anything 2 makes strong encoder for natural and medical image segmentation. *arXiv preprint arXiv:2408.08870*.
<https://doi.org/10.48550/arXiv.2408.08870>
- Yetiser, S., ve D. Ince. "Demographic analysis of benign paroxysmal positional vertigo as a common public health problem." *Annals of medical and health sciences research* 5.1 (2015): 50-53. <http://dx.doi.org/10.4103/2141-9248.149788>
- You, P., Instrum, R., & Parnes, L. (2019). Benign paroxysmal positional vertigo. *Laryngoscope investigative otolaryngology*, 4(1), 116-123.
<https://doi.org/10.1002/lio2.230>
- Zareapoor, M., Shamsolmoali, P., & Yang, J. (2019). Learning depth super-resolution by using multi-scale convolutional neural network. *Journal of Intelligent & Fuzzy Systems*, 36(2), 1773-1783. <https://doi.org/10.3233/JIFS-18136>
- Zhang, W., Wu, H., Liu, Y., Zheng, S., Liu, Z., Li, Y., ... & Zhu, Z. (2021). Deep learning based torsional nystagmus detection for dizziness and vertigo diagnosis. *Biomedical Signal Processing and Control*, 68, 102616.
<https://doi.org/10.1016/j.bspc.2021.102616>
- Zhou, Y., Sun, G., Li, Y., Xie, G. S., Benini, L., & Konukoglu, E. (2025). When sam2 meets video camouflaged object segmentation: A comprehensive evaluation and adaptation. *Visual Intelligence*, 3(1), 1-14.
<https://doi.org/10.48550/arXiv.2409.18653>

EKLER

EK 1: Etik Kurul Onayı

Evrak Tarih ve Sayısı: 14.10.2025-E.202535443



T.C.
İSTANBUL SABAHATTİN ZAİM ÜNİVERSİTESİ REKTÖRLÜĞÜ
ETİK KURULU BAŞKANLIĞI

Sayı : E-20292139-050.03-202535443
Konu : Etik Kurul 2025/9 No'lu Toplantı
(Hasibe Büşra AYTEKİN)

Sayın Arş. Gör. Hasibe Büşra AYTEKİN

"Açıklanabilir Yapay Zeka ile Benign Paroksizmal Pozisyonel Vertigo Tanısında Karar Destek Sisteminin Geliştirilmesi" başlıklı araştırmanızla ilgili başvurunuz, kurulumuzun 02.10.2025 tarihli ve 2025/9 sayılı toplantısında değerlendirilerek araştırmanız etik açıdan uygun bulunduğu katılanların oy birliği ile karar verilmiştir.

Bilgilerinize arz ederim.

Prof. Dr. Ali Haydar ŞAR
Kurul Başkanı

Ek:(Hasibe Busra AYTEKİN)Etik Kurul 2025_9 No_lu Toplantı (1 Sayfa)

Mevcut Elektronik İmzalar

ALİ HAYDAR ŞAR (Etik Kurulu Başkanlığı - Kurul Başkanı) 14.10.2025 15:05

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu : *BSL3LMS0Y1* Pin Kodu : 25762

Belge Takip Adresi : https://cbys.izu.edu.tr/enVision/Validate_Doc.aspx

Adres: Halkalı Caddesi No: 281 Küçükçekmece/İstanbul
Telefon: 444 97 98 Faks: +90 (212) 693 82 29
e-Posta: bilgi@izu.edu.tr Web: www.izu.edu.tr
Kep Adresi: izu@hs01.kep.tr

Bilgi için: Murat AKTAŞ
Unvanı: Yeminli Katip



 Istanbul Zaim Üniversitesi	ETİK ONAY BELGESİ	Doküman No	ET-FR-675
		İlk Yayın Tarihi	27.11.2023
		Revizyon Tarihi	22.12.2023
		Revizyon No	01
		Sayfa	1/1

Tarih	02.10.2025
Sayı	2025/9
Araştırmanın Niteliği	Doktora Tezi
Araştırmanın Adı	<i>Açıklanabilir Yapay Zeka ile Benign Paroksizmal Pozisyonel Vertigo Tamsında Karar Destek Sisteminin Geliştirilmesi</i>
Sorumlu/Yürütücü Araştırmacının Adı Soyadı	Hasibe Büşra AYTEKİN
Diğer Araştırmacı/lar Adı Soyadı	Prof. Dr. Tark ÖZKUL
Karar	UYGUNDUR

(İstanbul Sabahattin Zaim Üniversitesi Etik Kurulu'nun kararı tavsiye niteliğinde olup, Üniversitemizle ilgili etik ilkelerinin belirlenmesi ve değerlendirilmesi amacıyla taşımaktadır.)

(Katıldı)
Prof. Dr. Ali Haydar ŞAR
Başkan

(Katıldı)
Prof. Dr. Mehmet Emin KÖKTAŞ
Başkan V.

(Katılmadı)
Prof. Dr. Mustafa ATEŞ
Üye

(Katıldı)
Prof. Dr. Beytullah KAYA
Üye

(Katılmadı)
Prof. Dr. Ayşe Nefise BAHÇECİK
Üye

(Katılmadı)
Prof. Dr. Tark ÖZKUL
Üye

(Katıldı)
Av. Bilal ŞAMAT
Üye

Mevcut Elektronik İmzalar

ALİ HAYDAR ŞAR (Etik Kurulu Başkanlığı - Kurul Başkanı) 14.10.2025 15:05
Kurul Yeminli Kâtibi: Murat AKTAŞ

Dokümantasyon Onayı Prof. Dr. İbrahim GÜNEY Yönetim Sistemleri Temsilcisi

Bu belge, güvenli elektronik imza ile imzalanmıştır.
Evrak sorgulaması <https://ebys.izu.edu.tr/enVision/Sorgula/BSF3LFYY5Y> adresinden yapılabilir.

ÖZGEÇMİŞ

A. EĞİTİM BİLGİLERİ

Hasibe Büşra AYTEKİN, 2017 yılında İstanbul Sabahattin Zaim Üniversitesi Bilgisayar Mühendisliği lisans programından mezun olmuştur. 2020 yılında aynı üniversitenin Bilgisayar Bilimi ve Mühendisliği (%30 İngilizce) yüksek lisans programını tamamlamıştır. 2020 yılında İstanbul Sabahattin Zaim Üniversitesi Bilgisayar Bilimleri ve Mühendisliği (%30 İngilizce) doktora programına başlamıştır.

B. MESLEKİ DENEYİM

Araştırma Görevlisi, İstanbul Sabahattin Zaim Üniversitesi, Bilgisayar Mühendisliği Bölümü (2020 – devam)

C. YAYINLAR

Dogru, H. B., Kaplan, A., Zengin, A. T., & Ozkul, T. (2024, April). Image-based Nystagmus Analysis in BPPV Disorders: Polar Coordinate System and Template Matching. In *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)* (pp. 1-5). IEEE.

Dogru, H. B., & Zengin, A. T. (2022, June). Fast-ICA Based Lane Detection Method for Autonomous Vehicles. In *2022 26th International Conference Electronics* (pp. 1-6). IEEE.

Doğru, H. B., Şirin, Y., Tilki, S., Yeşiltepe, M., Rasheed, J., & Khan, M. U. S. (2021). Fruits Sorting with Instance Based Image Processing. In *1st International Conference on Computing and Machine Intelligence*. İstanbul Sabahattin Zaim Üniversitesi

- Tilki, S., Dogru, H. B., & Hameed, A. A. (2021). Gender classification using deep learning techniques. *Manchester journal of Artificial Intelligence and Applied sciences*, 2(2).
- Dogru, H. B., Tilki, S., Jamil, A., & Hameed, A. A. (2021, April). Deep learning-based classification of news texts using doc2vec model. In *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)* (pp. 91-96). IEEE.
- Dođru, H. B., Hameed, A. A., Tilki, S., & Jamil, A. (2021). Comparative analysis of deep learning and traditional machine learning models for Turkish text classification. In *1st International Conference on Computing and Machine Intelligence*. İstanbul Sabahattin Zaim Üniversitesi.
- Rasheed, J., Dogru, H. B., & Jamil, A. (2020, August). Turkish text detection system from videos using machine learning and deep learning techniques. In *2020 IEEE Third international conference on data stream mining & processing (DSMP)* (pp. 116-120). IEEE.
- Dođru, H. B. (2020). *Yapay zeka ile ortak yönelimli web sitelerinin tespiti* (Master's thesis, İstanbul Sabahattin Zaim Üniversitesi).
- Rasheed, J., Jamil, A., Dogru, H. B., Tilki, S., & Yesiltepe, M. (2019, November). A deep learning-based method for Turkish text detection from videos. In *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)* (pp. 935-939). IEEE.

D. PROJELER

Proje Adı: Benign Paroksizmal Pozisyonel Vertigo (BPPV) Bozuklukları Teşhisi için Görüntü Tabanlı Derin Öğrenme Sistemi

Proje No: BAP-2023-16

Destekleyen Kurum: İstanbul Sabahattin Zaim Üniversitesi

Proje Süresi: 2023 – 2025