

**T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DERİN ÖĞRENME TABANLI VE GERÇEK ZAMANLI BAKIŞ TAKİBİ

DOKTORA TEZİ

Kenan DONUK

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Davut HANBAY

MAYIS 2023

T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DERİN ÖĞRENME TABANLI VE GERÇEK ZAMANLI BAKIŞ TAKİBİ

DOKTORA TEZİ

Kenan DONUK
(36173619015)

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Davut HANBAY

MAYIS 2023

TEŞEKKÜR VE ÖNSÖZ

Bu tez çalışmasının ve doktora eğitimimin her aşamasında bana destek olan akademik bilgi ve tecrübelerinin yanı sıra yardım, öneri, bilgi ve desteklerini sabır ve cömertlikle benimle paylaşan kıymetli danışman hocam Prof. Dr. Davut HANBAY'a,

Tez çalışmam süresince değerli görüşleriyle tezime katkı sağlayan değerli Prof. Dr. Ali KARCI, Prof. Dr. Muhammed Fatih TALU, Dr. Öğr. Üyesi Ali ARI hocalarıma ve ayrıca bana destek olan Bilgisayar Mühendisliği Bölümünde görev yapan tüm diğer değerli hocalarıma,

Bu çalışma İnönü Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi (BAP) tarafından FDK-2020-2110 kodlu proje ile desteklendiğinden BAP birimine,

Hayatımın her anında bana destek olan, bugünlere gelmem de büyük emekleri olan, haklarını hiçbir zaman ödeyemeyeceğim ve varlığına her daim şükrettiğim kıymetli anneme, babama ve kardeşlerime,

Her konuda olduğu gibi bu süreçte de desteklerini esirgemeyen değerli eşim Hülya DONUK'a ve beni sabırla bekleyen çocuklarım Ahsen Beyza, Muhammed Furkan ve Rüya'ya

Sonsuz teşekkürlerimi sunarım.

ONUR SÖZÜ

Doktora tezi olarak sunduđum “Derin Öğrenme Tabanlı ve Gerçek Zamanlı Bakış Takibi” başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığına ve yararlandığım bütün kaynakların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuđunu belirtir, bunu onurumla doğrularım.

Kenan DONUK

İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ	i
ONUR SÖZÜ	ii
İÇİNDEKİLER.....	iii
ÇİZELGELER DİZİNİ.....	v
ŞEKİLLER DİZİNİ.....	vi
SEMBOLLER VE KISALTMALAR	viii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ.....	1
1.1 İnsan Gözünün Anatomisi	2
1.2 Bakış Takip Teknolojisi.....	4
1.3 Bakış Takip Teknikleri	4
1.3.1 Electro-oculografi yöntemi.....	5
1.3.2 Skleral kontakt lens yöntemi	5
1.3.3 Video tabanlı yöntemler	6
1.4 Tezin Amaç ve Hedefleri.....	10
1.5 Tezin Organizasyonu ve Yapılan Bilimsel Katkıları	10
2. YÜZ İMGELERİNDEN GÖZ BÖLGELERİNİN TESPİTİNDE ESA TABANLI ALTERNATİF BİR YAKLAŞIM.....	12
2.1 İlgili Çalışmalar	13
2.2 Materyal ve Metodoloji	14
2.2.1 Evrişimsel Sinir Ağları	14
2.2.2 Veri seti oluşturma	16
2.2.3 Önerilen ESA tabanlı sistem ve akış şeması	18
2.2.4 Sistem parametreleri ve eğitim	21
2.3 Deneysel Sonuçlar	22
3. MİNİ U-NET MİMARİSİNE DAYALI GÖZ BEBEĞİ MERKEZİ LOKALİZASYONU.....	26
3.1. İlgili Çalışmalar	27
3.2 U-Net Mimarisi.....	28
3.3 Materyal ve Metodoloji	29
3.3.1 Veri seti	29
3.3.2 Önerilen Mini U-Net mimarisi ve eğitimi	30
3.4 Deneysel Sonuçlar	33
4. GÖRÜNTÜ İŞLEMENE DAYALI GERÇEK ZAMANLI BAKIŞ TAKİBİ 37	37
4.1 İlgili Çalışmalar	37
4.2 Bakış Takip Sistemi ve Algoritma.....	38
4.2.1 Yüz ve göz tespiti	39
4.2.2 İris ve iris merkezinin tespiti	41
4.2.2.1 Medyan filtreleme.....	41
4.2.2.2 Eşikleme	41
4.2.2.3 Morfolojik açma	42
4.2.3 Kalibrasyon prosedürü.....	44
4.2.4 Ekran bakış yeri tayini.....	44
4.3 Deneysel Sonuçlar	45
5. DERİN ÖĞRENMEYE DAYALI GERÇEK ZAMANLI BAKIŞ TAKİBİ ..47	47
5.1 İlgili Çalışmalar	47
5.2 Materyal ve Metodoloji	50
5.2.1 Veri seti oluşturma	50

5.2.2 Yüz dönüm noktaları	51
5.2.3 Veri seti temizleme ve düzenleme.....	52
5.2.4 Sistem ESA modeli ve eğitim.....	53
5.3 Deneysel Sonuçlar	57
6. BAKIŞ TAKİBİNİN ÇOCUKLARDA DİKKAT EKSİKLİĞİ TANISINDA KULLANIMI	60
6.1 İlgili Çalışmalar	60
6.2 Materyal ve Metodoloji	61
6.2.1 Katılımcılar.....	61
6.2.2 Cihazlar ve teknik özellikleri.....	62
6.2.3 Deney tasarımı ve ölçüm metrikleri	63
6.2.4 Destek Vektör Makinesi ve eğitimi	66
6.3 Deneysel Sonuçlar	67
7. SONUÇLAR.....	70
KAYNAKLAR.....	71
ÖZGEÇMİŞ	83

ÇİZELGELER DİZİNİ

Çizelge 2.1: Sistem ESA modeli verileri	21
Çizelge 2.2: 600 görüntü üzerinde karşılaştırma.....	23
Çizelge 3.1: GI4E veri seti performans karşılaştırması	35
Çizelge 4.1: Görüntü işleme adımları	42
Çizelge 5.1: Sistem ESA modeli eğitim ve test verileri.....	57
Çizelge 6.1: GP3 HD bakış takip cihazı teknik özellikleri	62
Çizelge 6.2: Katılımcılara ait ölçüm metrik değerleri.....	65
Çizelge 6.3: Hasta ve sağlıklılara ait ölçüm ortalamaları	66

ŞEKİLLER DİZİNİ

Şekil 1.1: İnsan gözünün yapısı	3
Şekil 1.2: Optik ve görsel eksen.....	3
Şekil 1.3: EOG yönteminin uygulanması	5
Şekil 1.4: SKL yönteminin uygulanması	6
Şekil 1.5: (a) Gözdeki yansıma (b) Göz kapaklarının durumu	6
Şekil 1.6: Sabit baş pozisyonunda göz hareketleri.....	7
Şekil 1.7: Hareketli baş pozisyonunda göz hareketleri	7
Şekil 1.8: GKBY ilişkin retina yansıması	8
Şekil 1.9: GKBY ilişkin kornea yansıması	8
Şekil 1.10: Video tabanlı bakış takip sistemlerinin genel işleyiş şeması.....	9
Şekil 2.1: LeNet-5 ESA yapısı	15
Şekil 2.2: Kırpılmış göz görüntüleri.....	16
Şekil 2.3: Yüzde göz olmayan bölgeler	17
Şekil 2.4: Kapalı göz görüntüleri	17
Şekil 2.5: Rastgele parlaklık	18
Şekil 2.6: Rastgele rotasyon	18
Şekil 2.7: Önerilen ESA tabanlı sistem.....	19
Şekil 2.8: Önerilen sistemin ESA modeli	19
Şekil 2.9: 1 nolu evrişim katmanı çıktıları	20
Şekil 2.10: Maksimum havuzlama katmanı çıktısı	20
Şekil 2.11: Model doğruluk grafiği	21
Şekil 2.12: Model kayıp grafiği	22
Şekil 2.13: Viola-Jones algoritması uygulanan görüntüler	22
Şekil 2.14: Önerilen ESA tabanlı sistemin uygulandığı görüntüler.....	23
Şekil 2.15: Viola-Jones ve Önerilen sistem doğruluk karşılaştırması.....	24
Şekil 3.1: U-Net mimarisi	29
Şekil 3.2: GI4E veri seti görüntüleri	30
Şekil 3.3: GI4E veri seti görüntüsüne yer işaretlerinin uygulanması.....	30
Şekil 3.4: Mini U-Net ağ yapısı	31
Şekil 3.5: Göz görüntülerinin maskelenmesi	32
Şekil 3.6: Test verisi lokalizasyon tahminleri	33
Şekil 3.7: Normalleştirilmiş hata değeri sınırları	34
Şekil 3.8: Gerçek ve tahmin lokalizasyonlarının karşılaştırılması.....	34
Şekil 3.9: Gerçek zamanlı gözbebeği lokalizasyonu.....	36
Şekil 4.1: Bakış takip sistem kurulumu	38
Şekil 4.2: Bakış takibi sistemi akış şeması	39
Şekil 4.3: 68 adet dönüm noktası gösterimi	40
Şekil 4.4: Dönüm noktalarının yüze uygulanması	41
Şekil 4.5: Kontur uygulaması ve ağırlık merkezi.....	43
Şekil 4.6: Kalibrasyon prosedürü	44
Şekil 4.7: Bakış yeri doğrulama	46
Şekil 5.1: Veri toplama uygulama ekranı.....	50
Şekil 5.2: Dönüm noktalarının yüze uygulanması	51
Şekil 5.3: (a) Yüz, (b) Yüz konumu, (c) Sol göz, (d) Sağ göz, (e) Baş pozisyon bilgisi	52
Şekil 5.4: Kapalı göz tespiti	53
Şekil 5.5: Tasarlanan ESA tabanlı sistem	54
Şekil 5.6: Öklit mesafesi	55

Şekil 5.7: Model doğruluk grafiği	56
Şekil 5.8: Model kayıp grafiği	56
Şekil 5.9: Odaklanma doğruluk testi	57
Şekil 5.10: Bilgi arama deneyi ve ısı haritası gösterimi	58
Şekil 6.1: GP3 HD bakış takip cihazı	62
Şekil 6.2: Kalibrasyon ekranı	63
Şekil 6.3: Katılımcı ve cihaz düzeni	63
Şekil 6.4: Görev topu senaryosu	64
Şekil 6.5: DVM tabanlı sistem	66
Şekil 6.6: Performans metrikleri	68
Şekil 6.7: Karışıklık matrisleri	68

SEMBOLLER VE KISALTMALAR

EOG	: Electro-OculoGraphy
SKL	: Skleral Kontakt Lens
VTY	: Video Tabanlı Yöntem
GBKY	: Göz Bebeği Merkezi-Kornea Yansıması
YSA	: Yapay Sinir Ağı
ESA	: Evrişimsel Sinir Ağı
EYK	: En Yakın Komşu Algoritması
GS	: Gauss Süreci
DVM	: Destek Vektör Makineleri
RO	: Ratgele Orman
Dİ	: Doğrusal İnterpolasyon
GBESA	: Çok Görevli Basamaklı Evrişimsel Sinir Ağı
VF	: Varyans Filtresi
DSA	: Derin Sinir Ağı
UKSB	: Uzun-Kısa Süreli Bellek
GEBO	: Göz En-Boy Oranı
DBM	: Derin Boltzman Makineleri
TSA	: Tekrarlayan Sinir Ağı
TEA	: Tam Evrişimsel Ağ
OKH	: Ortalama Kare Hata
UDR	: Uyarlanabilir Doğrusal Regresyon
SSGM	: Sağ-Sol Göz Arası Mesafe
d	: Piksel cinsinden mesafe
f	: Filtre
p	: Dolgu
F_c	: Evrişim çıktısı
I	: Girdi görüntüsü
K	: Çekirdek
2B	: İki boyut
3B	: Üç boyut
İA	: İlgili alanı
İAD	: İlgili alanı dışı
TZ	: Tekrar ziyaret

ÖZET

Doktora Tezi

DERİN ÖĞRENME TABANLI VE GERÇEK ZAMANLI BAKIŞ TAKİBİ

KENAN DONUK

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

83+X sayfa

2023

Danışman: Prof. Dr. Davut HANBAY

Bu tez çalışması yüksek bütçeli bakış takip cihazlarına alternatif derin öğrenme tabanlı ve düşük maliyetli sistemlerin geliştirilebileceğini, bakış takibi için önemli girdiler olan hem göz hem de gözbebeği merkezinin konumlarını hassas bir şekilde tespit etmek için Evrişimsel Sinir Ağı tabanlı teknikler kullanılarak yüksek düzeyde doğruluk elde edilebileceğini gösteren yöntemler sunmaktadır. Ayrıca çalışmada, çocuklarda dikkat eksikliği teşhisinde bakış takip teknolojisi yoluyla yakalanan zengin verilerden bireylerin dikkat kalıplarına ilişkin değerli iç görüler elde edilebileceğini gösteren klinisyenlere yardımcı bir karar destek sistemi geliştirilmiştir. Çalışmada önerilen ilk yöntem olan göz tespiti, bakış takip teknolojisinin çalışma prensibinde ilk aşamadır. Bakış takibinin doğru bir şekilde yapılması sağlam bir göz tespitine dayanır. Bunun için Evrişimsel Sinir Ağlarına dayanan bir göz algılama yöntemi önerilmiştir. İkinci yöntem, görme olayında aktif rol oynayan göz bebeği merkezinin yüksek doğrulukta tespitine dayanmaktadır. Göz bebeği merkezinin sağlam bir şekilde tespit edilebilmesi bakış takip doğruluğunu önemli ölçüde etkilemektedir. Bu nedenle göz bebeği merkezi tespitinde Evrişimsel Sinir Ağlarına dayalı bir Mini U-Net yapısı önerilmiştir. Üçüncü ve dördüncü yöntem sırasıyla bakış takibinin görüntü işlemeye ve Evrişimsel Sinir Ağlarına dayalı, düşük maliyetli uygulamalarıdır. Tezin son yönteminde günümüz bakış takip teknolojisi ile makine öğrenmesi birleştirilerek bireylerde dikkat eksikliği tanısının konulmasında klinisyenlere yardımcı bir karar destek sistemi önerilmiştir. Bu tez çalışması ayrıntılı ve tanımlayıcı metodolojiler sağlayarak, bakış takibini çeşitli alanlarda daha erişilebilir ve anlamlı hale getirmeyi amaçlamaktadır.

Anahtar Kelimeler: Göz Tespiti, Göz Bebeği Merkezi Lokalizasyonu, Göz Takibi, Dehb

ABSTRACT

Phd. Thesis

DEEP LEARNING BASED AND REAL-TIME GAZE TRACKING

Kenan DONUK

Inonu University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

83+X page

2023

Supervisor: Prof. Dr. Davut HANBAY

This thesis work shows that alternative Deep Learning-based and low-cost systems can be developed for high-priced gaze-tracking devices and that high accuracy can be achieved by using Convolutional Neural Network-based techniques to accurately capture the positions of the eye and pupil center, which are important inputs for eye-tracking. provides. In addition, the study developed a decision support system for clinicians, demonstrating that valuable insights into individual attention patterns can be gained from the rich data captured by gaze-tracking technology in diagnosing attention deficits in children. Eye detection, the first method proposed in the study, is the first step in the working principle of gaze-tracking technology. Accurate gaze tracking depends on solid eye recognition. For this purpose, an eye detection method based on convolutional neural networks is proposed. The second method is based on high-precision detection of the pupil center, which plays an active role in vision. The ability to reliably detect the pupil center has a significant impact on the accuracy of gaze tracking. Therefore, a Mini-UNet network structure based on convolutional neural networks has been proposed for pupil center detection. The third and fourth methods are low-cost applications of gaze tracking based on image processing and convolutional neural networks, respectively. In the last method of the work, a decision support system is proposed to help clinicians diagnose attention deficits in individuals by combining current eye-tracking technology and machine learning. This thesis work aims to make gaze-tracking more accessible and useful in various domains by providing detailed and descriptive methods.

Keywords: Eye Detection, Pupil Center Localization, Gaze Tracking, Adhd

1. GİRİŞ

Göz, insanda görme işleminin gerçekleştiği bir duyu organıdır. İnsan-çevre ilişkisindeki duyuusal bağlantının büyük bir bölümü göz aracılığıyla gerçekleşmekte ve insan beyninin bilişsel süreçlerini tetikleyen bir pencere görevi görmektedir. Gözün insan hayatındaki bu önemli rolü bakış takip çalışmalarına ilham kaynağı olmuştur. Bakış takip teknolojisi, hayatın hemen hemen her alanında kullanılabilen teknolojik ürünlerin kullanımını kolaylaştıran ve etkileşim verimliliğini arttıran temassız bir kontrol yöntemidir. Beyne iletilen bilgilerin %90'ının gözler tarafından algılandığı düşünüldüğünde bu teknolojinin önemi daha iyi anlaşılmaktadır [1]. Tarihi 19. yüzyıla dayanan bakış takip çalışmaları [2] ilk olarak gözdeki kas kasılmalarının çıkardıkları sesler ve göz hareketlerinin fotografik kayıtları elde edilerek bazı araştırmalar yapılmıştır. Fakat bu araştırmalarda göz hareketlerinin takip ve kontrolünün yüksek doğrulukta yapılması pek mümkün olmamıştır. 20. yüzyıla girildiğinde gelişen bakış takip araçlarıyla birlikte bu teknolojiye ilerleme sağlanmıştır.

Günümüzde gelişen teknolojiyle beraber her alanda olduğu gibi insan-bilgisayar etkileşimi araştırmalarında da önemli gelişmeler yaşanmaktadır. İnsan-Bilgisayar etkileşimi alanındaki en aktif araştırmalardan biri de bakış takip teknolojisidir. Bu teknoloji kişinin nereye, ne kadar süre baktığını ölçen ve bu ölçümlerden yola çıkarak kişinin ilgi alanları, bilişsel durumu, duygusal durumları ve niyeti hakkında tahminler yürütebilmesini sağlayan bir sensör teknolojisidir. Dünyaca ünlü Grand Vime Research danışmanlık şirketinin 2022 yılı raporuna göre bakış takip pazarının 2022 yılı pazar büyüklüğü 806,8 milyon ABD doları ve 2030 yılına kadar küresel bazda 8,06 milyar ABD dolarına ulaşması öngörülmüştür [3]. Bu öngörüler ışığında Facebook gibi bazı büyük firmalar bu alanda yatırımları söz konusudur [4]. Bu teknoloji sayesinde göz hareketlerinden elde edilen verilerin analiz edilmesiyle tıbbi araştırmaların çeşitli nörolojik-psikiyatrik durumlarının incelemesinde, eğitim ve öğrenme metotlarında, web sayfalarının tasarımında, reklam seçimlerinde, pazar araştırmalarında, oyun-eğlence sektöründe, otomotivde sürüş güvenliğinde, fiziksel engelli bireylerde iletişim kurma ve çocuklarda dikkat eksikliği [5–13] gibi birçok farklı alanda göz hareket verileri etkin olarak kullanılmaktadır.

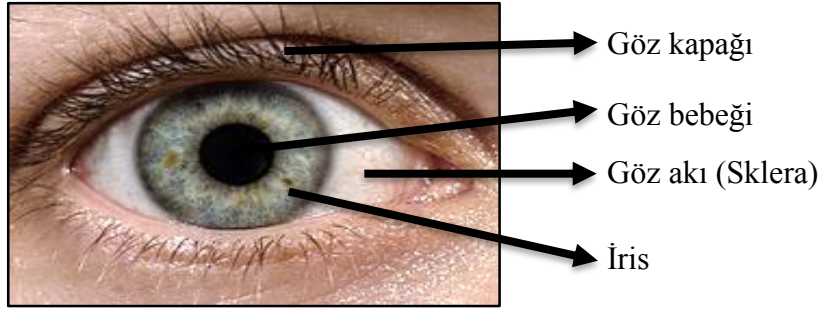
Bakış takibinde geçmişten günümüze birçok farklı yöntem uygulanmıştır. Uygulanan bu yöntemler sağlık ve doğruluk açısından değerlendirilmektedir. Günümüzde yaygın olarak kızılötesi tabanlı yöntemlere dayalı ticari bakış takip cihazları kullanılmaktadır. Bu cihazlar bir veya daha fazla yüksek çözünürlüklü kamera, kızılötesi ışık kaynakları ve bilgisayarlı görme algoritmalarından oluşur. Bu cihazlarda bakış takibi kameradan alınan bilgilerin makine öğrenimi, gelişmiş görüntü işleme süreçleri ve matematiksel hesaplamalar yardımıyla bakış koordinatlarına dönüştürülerek yapılır. Bu cihazlar bazı kısıtlamalar eşliğinde yüksek doğrulukta bakış tahminleri yapsa da sağlık açısından henüz yeterli değildirler. Ayrıca bu cihazlar yüksek maliyetli donanımlara sahip olduğundan kullanım alanları araştırma-geliştirme alanlarıyla sınırlıdır.

Bakış takip teknolojisi, uzun yıllardır var olmasına rağmen tüketici uygulamalarında etkin bir kullanım alanına sahip değildir. Bu eksikliğin sebebi bakış takip donanımının yüksek maliyeti, her ortamda kullanımının zorluğu ve tüketici düzeyinde uygulamaların eksikliğidir [14]. Bakış takibinin herkes tarafından deneyimlenmesi için tüketici sınıfı akıllı telefon, dizüstü bilgisayar ve tabletlerde yerleşik olarak kullanılacak, daha düşük çözünürlükte çalışabilen, değişen ortam koşullarına karşı sağlam bir yöntem ihtiyacı vardır. Son yıllarda bu ihtiyaçları karşılayacak üzerinde yoğun çalışmaların yapıldığı aktif bir çalışma alanı olan derin öğrenme tabanlı bakış takip yöntemleri geliştirilmektedir.

1.1 İnsan Gözünün Anatomisi

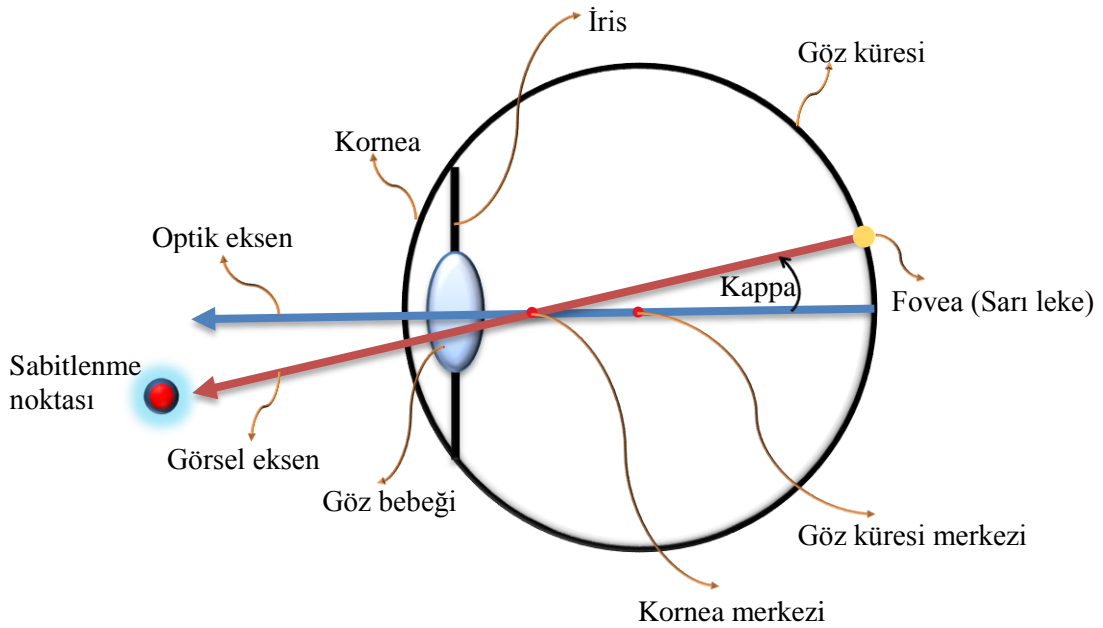
İnsan vücudunun karmaşık organlarından biri olan gözler, göz çukurunda bulunan, küre şeklindeki yapısıyla görmeyi sağlayan bir organdır. Gözün görünen kısmı sklera, iris, göz bebeği ve göz kapağı ile dört kısımdan oluşur. Sklera gözün dış bölgesini oluşturur. Sklera, gözü iç ve dış kuvvetlerden koruyan bağ dokusuna sahip sert tabakayı oluşturur. Sklera ya da “Gözakı” gözün ön tarafında tümsek oluşturarak, saydam kornea tabakasını oluşturur. Gözün orta tabakasını oluşturan ve rengi insandan insana değişen iris, ön bölümde damar tabaka (koroid)’nin renk değiştirmesiyle ortasında göz bebeğinin olduğu bir diyafram görevi görür. İrisin sahip olduğu ve dikey olarak konumlanan kas lifleri kasılarak gözbebeğinin genişlemesini, irisin etrafında çember şeklinde konumlanan kas lifleri ise, kasıldıklarında göz bebeğinin büzülmesini sağlar. Sonuç olarak iris, değişen ortam koşullarına göre, gözün içine giren ışık miktarını ayarlayan bir diyafram görevi görerek göz bebeğinin çapını belirler [15,16]. Kornea göze giren ışığı kırarak merceğe ve retinaya iletir. Son olarak retinanın merkezi olarak kabul edilen fovea (sarı leke) üzerinde ters bir görüntü

oluşturur. Bu sürecin sonun görme olayı gerçekleşir. Şekil 1.1’de insan gözünün bazı kısımları gösterilmiştir.



Şekil 1.1: İnsan gözünün yapısı [17]

Gözün görme sürecinde literatürde bilinen iki önemli eksen vardır. Bunlardan ilki optik eksenidir. Bu eksen kornea eğrilik merkezi ile göz küresinin eğrilik merkezini birleştiren sanal bir hattır. Diğer eksen ise görüş hattı olarak da adlandırılan görsel eksenidir. Görsel eksen, görme olayının gerçekleştiği retinanın merkezi olarak bilinen ve görüntüye ait en net ve en detaylı bilginin algılandığı fovea ile kornea eğrilik merkezini birleştiren sanal hattır. Dolayısıyla görsel dikkati veya bakış yönünü belirleyen optik eksen değil, görsel eksenidir. Optik eksen ile görsel eksen arasında kesişimden dolayı 5°’lik (Kappa açısı) bir açı mevcuttur [18,19]. Kappa açısı kişiye özel bir açı değerine sahiptir. Şekil 1.2’de insan gözünün optik ve görsel eksenini gösterilmiştir.



Şekil 1.2: Optik ve görsel eksen

1.2 Bakış Takip Teknolojisi

Bakış takip teknolojisi göz hareketlerini takip eden ve gözün nereye baktığını gerçek zamanlı olarak tespit eden bir sensör teknolojisidir [14]. Bu teknolojinin ticari kullanımında genel olarak giyilebilir ve ekran tabanlı cihazları mevcuttur. Gelişmiş donanım ve algoritmalar kullanarak bakış takibi gerçekleştiren cihazlar üreten ve bakış takibi için veri analizi yazılımları üreten SMI, EyeLink, Smart Eye, Gazepoint, Tobii, Imotions, EyeTech gibi farklı firmalar mevcuttur [20]. Bakış takip teknolojisinin gelişiminde geçmişten günümüze birçok farklı teknik kullanılmıştır. Bu teknikler sağlamlık ve doğruluk açısından değerlendirilmektedir. En yaygın kullanılan teknik, kızılötesi ışığın ve yüksek çözünürlüklü kameraların kullanıldığı tekniktir. Bu tekniğin çalışma prensibi, göz bebeği merkezi-kornea yansımalarına dayanır. Bakış takip teknolojisinde elde edilen bilgilerle anlamlı veriler elde etmek için veri analizleri gerçekleştirilir. Veri analizi için farklı örneklem hızlarında (150Hz, 300Hz) çalışan cihazların elde ettikleri bazı önemli parametreler mevcuttur. Bunlar ilk fiksasyon süresi, fiksasyon süresi, ortalama fiksasyon süresi, fiksasyon sayısı, kalma süresi (dwell time), tekrar ziyaret sayısı (revisit), sakkad genliği, sakkad sayısı gibi ölçüm metrikleridir. Bunlardan en önemlileri fiksasyon ve sakkadlardır. Fiksasyonlar, gözün tarama işlemi esnasında belirli bir noktada veya alanda 50ms ile 600ms arasındaki bir sürede sabitlenmesi olarak ifade edilebilir [21]. Belirtilen süreler arasındaki bir fiksasyon veya sabitlenme görsel sistemin ilgili noktalardan veya alandan bilgi alımı gerçekleştirdiği diğer bir ifadeyle bilişsel bilgiler hakkında çıkarımın yapıldığı süreçlerdir. Sakkadlar ise gözün tarama esnasında bilgi alımı yapmadığı “atlama” olarak adlandırılan iki fiksasyon arasındaki gözün 100ms ile 1000ms süre aralığındaki anlık hareketidir [21–23]. Fiksasyon ve sakkad bilgilerinin verilere dönüştürülmesiyle bakış takibi süresince en çok dikkat çekilen yerler, izleme sırası-düzeni-süresi, en yüksek gözlemin yapıldığı bölgeler diğer bir ifadeyle ilgi alanları bakış takip cihazı yazılımları ile elde edilmektedir [24].

1.3 Bakış Takip Teknikleri

Bakış takibinde geçmişten günümüze birçok yöntem kullanılmıştır. Bu yöntemlerin bir kısmı invazif (müdahaleci) bir kısmı da invazif olmayan yöntemlerdir. Electro-OculoGraphy (EOG), Skleral Kontakt Lens (SKL) gibi invazif yöntemler daha çok erken dönem yöntemlerde karşımıza çıkmaktadır [2]. Bu yöntemlerde uygulamada kullanılacak ekipmanlar söz konusudur. Ayrıca bu yöntemler kullanıcının sabit baş pozisyonunu gerektirir.

Gelişen teknolojiyle beraber erken dönem bakış takip teknolojilerinden farklı olarak invazif olmayan video tabanlı yöntemler geniş çapta kullanım alanı bulmuştur. Bu yöntemin kullanıcı ile temas gerektirmemesi (kafaya giyilen kameralar hariç) ve kameradan alınan anlık bilgilerin analizine dayalı bir bakış takibi sunması ile bakış takip teknolojisine yeni bir boyut kazandırmıştır.

1.3.1 Electro-Oculografi yöntemi

Bakış takibinde kullanımı oldukça eski yöntemlerden biridir. Bu yöntem elektrot ekipmanları ile uygulanmaktadır. Göz çevresine yapıştırılmış elektrotlar kullanıldığından invazif bir yöntemdir. Elektro-Oculografi (EOG)'nin çalışma prensibi, göz çevresine yerleştirilen elektrotların kornea-retina arasında var olan elektriksel potansiyel farklılıklarının ölçülmesine dayanır [25]. Bu potansiyel, fundusta (gözün iç kısmı) negatif bir kutup ve korneada pozitif bir kutup bulunan sabit bir elektriksel dipol (kutuplaşma) olarak düşünülebilir. Böylece gözdeki hareket potansiyeli, göz merceğinin yeri değiştikçe gözlerin etrafına cilt üzerine yerleştirilmiş bir elektrot sistemi boyunca indüklenen voltajın ölçülmesi ile elde edilebilir [26]. Düşük hesaplama gücü gerektiren bu yöntemde bakış açısı ölçümü baş pozisyonuna göre gerçekleştirilir [25], [27]. Bu yöntemde büyük göz hareketlerinin yönü hasas bir şekilde elde edilebilirken küçük göz hareketleri için aynı hassasiyet söz konusu değildir. EOG kullanımına ilişkin görsel Şekil 1.3'te gösterilmiştir.

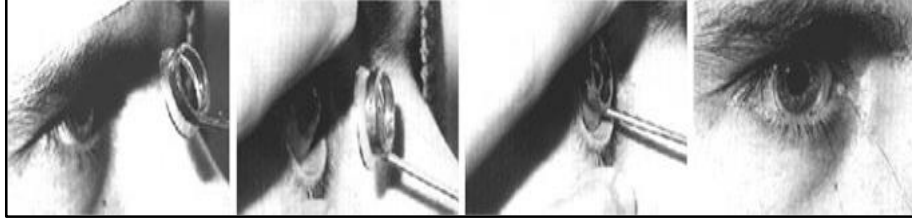


Şekil 1.3: EOG yönteminin uygulanması [25]

1.3.2 Skleral Kontakt Lens yöntemi

Bu yöntem özellikle tıbbi araştırmalarda kullanılmaktadır. Robinson tarafından 1963 yılında geliştirilen bu yöntem Collewijn ve arkadaşları tarafından iyileştirilmiştir [26,27]. Skleral Kontakt Lens (SKL) yönteminin çalışma prensibi, gözün sklerasına oturan bir silikon halkanın içine gömülü bir tel arama bobininden oluşmaktadır. Buradan ince bir kablo, harici bir ölçüm birimine bağlanır. Harici jeneratör bobinleri, bobinlerde yükseltilen ve ölçülen bir

voltajı indükleyen bir manyetik alan oluşturur [26–28]. Lensin takılmasından dolayı invazif (müdahaleci) bir yöntemdir. Bu yöntem ayrıca başa göre göz pozisyonunu da ölçer ve genel olarak bakış açısı ölçümü için uygun değildir [25]. Günümüzde bu yöntem üzerine hala araştırmalar yapılmakta ve geliştirilmektedir [29]. SKL yönteminin uygulanmasına ait görsel Şekil 1.4’te gösterilmiştir.



Şekil 1.4: SKL yönteminin uygulanması [25]

1.3.3 Video tabanlı yöntemler

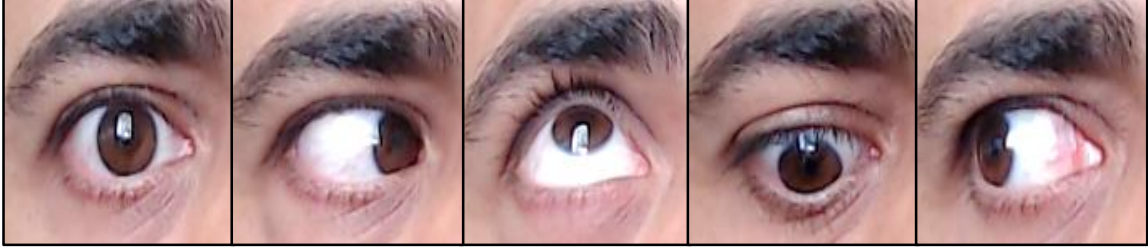
Video Tabanlı Yöntemler (VTY) yüksek veya düşük çözünürlüklü kamera veya kameraların kullanıldığı yöntemlerdir. VTY’lerde genel olarak şekil-model ve görünüm tabanlı bakış takibi yapan çalışmalar mevcuttur. Şekil-model tabanlı bakış tahmini çalışmalarında gözün bakış yönünü bulmak için göz bebeği merkezi, göz köşeleri, konturlar, kızılötesi ışık kaynağı kullanımı ve gözün değişmez geometrik özellikleri kullanır [30–33]. Bu yöntemler şekil-model tabanlı yöntemler olarak adlandırılmaktadır. Bu yöntemlerin genel dezavantajı gözün iris ve gözbebeği bölgelerinin tespitinin ortam ışığından kaynaklı gözdeki yansımalar ve kapalı göz kapakları yüzünden zor olmasıdır. Şekil 1.5’te ortam ışığından kaynaklanan göz yansıması ve kapalı göz kapaklarının durumu gösterilmiştir.



Şekil 1.5: (a) Gözdeki yansıma (b) Göz kapaklarının durumu

Bu yöntemlerin uygulanmasında genellikle görüntü ve ekran ilişkisinin kurulması ve her kullanıcı için kalibrasyon ve baş pozisyonu için bir kontrol prosedürü uygulanır. Her ne kadar doğru bakış tahmini için bakış ve ekran ilişkisinin kalibrasyonu yapılsa da bu ilişki baş pozisyonundan önemli ölçüde etkilenmektedir. Örneğin, görüş açısında veya baş

pozisyonundaki küçük bir deęişiklik sırasıyla Şekil 1.6 ve Şekil 1.7’de gösterildięi gibi gözün sırasıyla ileri, sola, yukarı, aşıęı ve saęa belirli hedeflere baktığı pozisyonlarda, göz görünümünde veya bakış yönünde önemli deęişikliklere neden olur [34].



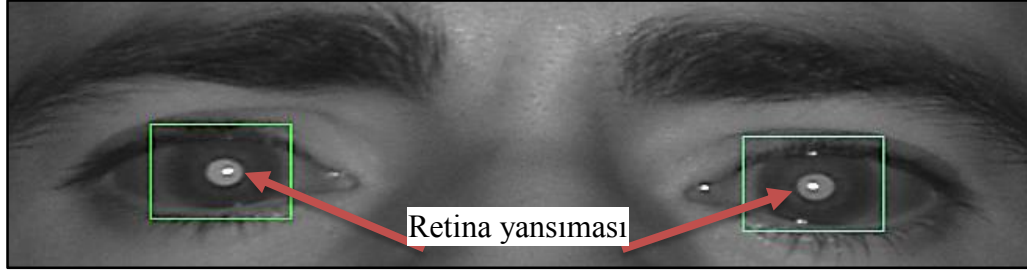
Şekil 1.6: Sabit baş pozisyonunda göz hareketleri



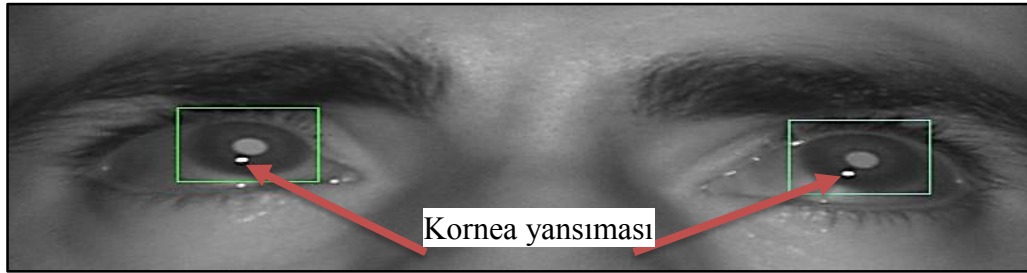
Şekil 1.7: Hareketli baş pozisyonunda göz hareketleri

Günümüz bakış takibinde kullanılan yöntemlerden en yaygın olanı kızılötesi ışık kaynağının kullanıldığı Göz Bebeęi-Kornea Yansıması (GBKY) yöntemidir. GBKY yaygın bir yöntem olmanın yanında ticari bakış takip cihazlarında da tercih edilmektedir. Bu yöntemde gözleri rahatsız etmeyen kızılötesi ışık kullanılmaktadır. Kızılötesi ışık sayesinde “kornea yansıması” veya “parıltı” olarak adlandırılan bir referans noktası kullanılır. Bu referans noktası, kamera eksenini üzerinde veya kamera eksenini dışında bir kızılötesi ışık kaynağı kullanılarak gerçekleştirilir. Kamera eksenini üzerindeki bir ışık kaynağı, “parlak gözbebeęi etkisi”, eksen dışındaki bir kızılötesi ışık kaynağı ise “karanlık göz bebeęi etkisi” ile sonuçlanmaktadır. Her iki durumda da renk yoğunluęu farkından dolayı göz bebeęinin ve merkezinin tespiti kolayca yapılabilmektedir. Göz bebeęinin merkezi ile kornea yansıması arasındaki vektörel ilişkidir yola çıkarak görüntüdeki gözbebeęi merkezinin göreceli konumuna göre bakış yönü tespit edilebilir [25,35]. Bu yöntemde kızılötesi aydınlatma kullanıldığından iç mekân uygulamalarında iyi sonuçlar verirken dış mekân uygulamalarında aydınlatma koşullardan dolayı kullanımı kısıtlıdır [36].

Bu yöntemde baş hareketine belirli ölçülerde izin verilmektedir. Fakat sabit baş pozisyonu doğru bakış tahmini için bu yöntemin önemli bir prosedürdür. Kızılötesi tabanlı ticari bakış takip cihazları mevcut son teknoloji yaklaşımlardan daha yüksek doğruluğa sahiptir [37]. Şekil 1.8 ve Şekil 1.9’da Gazepoint firmasına ait bakış takip cihazının yazılımından elde edilen GBKY’ye ilişkin görsel verilmiştir. Literatürde kızılötesi tabanlı bakış takibi üzerine yapılan birçok çalışma mevcuttur [38–40].

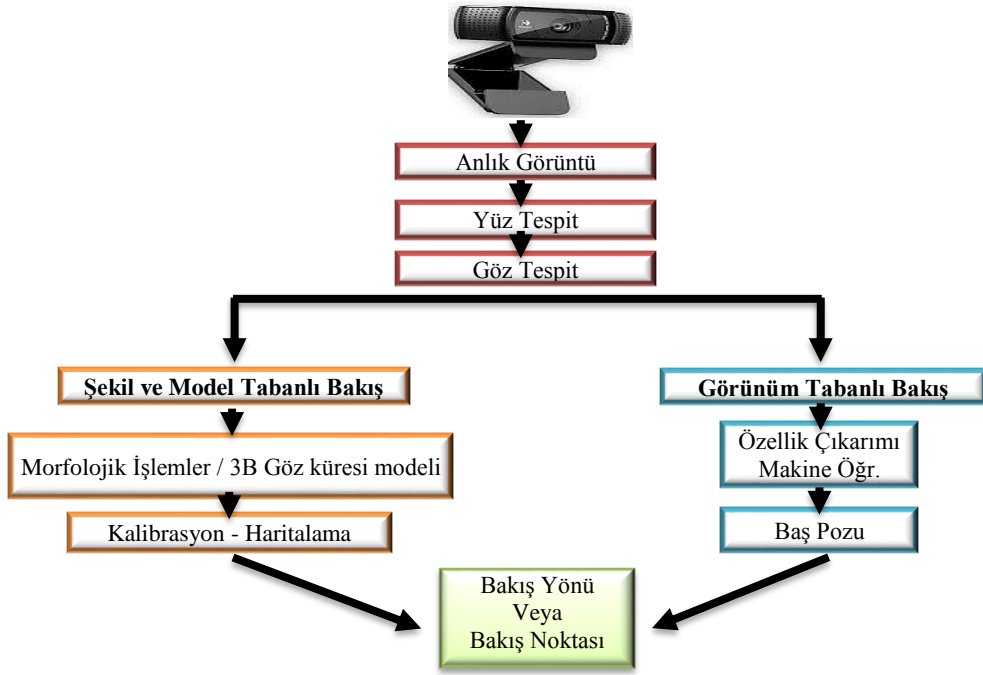


Şekil 1.8: GBKY ilişkili retina yansıması [41]



Şekil 1.9: GBKY ilişkili kornea yansıması [41]

Bakış takibinin yüksek doğrulukta ve sağlam bir şekilde yapılması için bakış doğruluğunu sınırlayan kalibrasyon ve baş hareketi problemlerinin çözümü için 3B model tabanlı yöntemler geliştirilmiştir. Bu yöntemlerde gözün konumu bakış tahmini için yeterli olmadığından buna ek olarak baş pozisyonu ve gözün yuvaları içindeki yönelimi de dikkate alınmıştır [42]. Dolayısıyla gözün konumuna ek olarak kişinin baş ve gözlerinin 3B görüş hattının da belirlenmesi ile bakış tahmini gerçekleştirilir [34]. 3B görüş hattının belirlenmesinde birden fazla kızılötesi ışık kaynağı veya kamera kullanılmaktadır. Bu sebeple maliyetli bir yöntemdir. VTY ile yapılan çalışmalarda gözün konumuna ek olarak bakış tahmini için göz küresi modeli [43–45] ve hareketli baş pozisyonu [33,46] modellerini kullanan çalışmalar yapılmıştır. Video tabanlı yöntemlerin genel işleyiş şeması Şekil 1.10’da gösterilmiştir.



Şekil 1.10: Video tabanlı bakış takip sistemlerinin genel işleyiş şeması

Son yıllarda gelişen teknolojiyle beraber makine öğrenimi yöntemleri bakış takip teknolojisinde geniş kullanım alanı bulmuştur. Geleneksel bakış takibi yapan diğer yöntemlerin aksine farklı ortam koşulları ve göz kapakları gibi bakış takibinde tıkanmaya sebep olan faktörlerin makine tarafından öğrenilmesiyle sağlam ve doğru bir bakış takibi sunmaktadır. Bu yöntem ile bakış tahminleri yapmak için farklı dış mekân koşullarında, farklı bireylerden fonksiyonel kameralar aracılığıyla alınan anlık göz ve yüz görüntülerinden oluşturulan veri kümelerine makine öğrenimi teknikleri uygulanır [37]. Makine öğrenimi ile görüntülerin düşük boyutlu bir öznitelik uzayı oluşturulur. Oluşturulan bu öznitelik uzayı görüntüleri temsil eder. Makine öğrenimi yöntemlerinden elde edilen öznitelikler ile bakış yönü arasındaki ilişki ile bakış takibi gerçekleştiren yöntemler görünüm tabanlı yöntemler olarak adlandırılmaktadır. Literatürde görünüm tabanlı bakış takibi için çeşitli algoritmalar kullanılmıştır. Bu algoritmalarından bazıları; Yapay Sinir Ağları (YSA) [47], Evrimsel Sinir Ağları (ESA) [48], En Yakın Komşu (EYK) [49], Gauss Süreci (GS) [50], Destek Vektör Makineleri (DVM) [51], Rastgele Orman (RO) [52] ve Doğrusal İnterpolasyon (Dİ) algoritmasıdır [53]. Görünüm tabanlı yöntemler yüksek maliyetli donanımlar, kalibrasyon ve kısıtlayıcı deney ortamları gerektirmediğinden yaygın olarak kullanılmasının önünü açmaktadır.

1.4 Tezin Amaç ve Hedefleri

Bu tez çalışması son teknoloji bakış takip cihazlarına alternatif düşük maliyetli çözüm yöntemleri ve bakış takip cihazlarının çocuklarda dikkat eksikliğinin tanısında kullanılabileceğini gösteren yöntemler ve uygulamalar sunmaktadır.

İlk olarak bakış takip sistemlerinde bakış yönünün tespitinin önemli bir parçası olan göz tespiti ve göz bebeği merkezi lokalizasyonu için derin öğrenme tabanlı iki ayrı yöntem önerilmiştir. Bu yöntemler bakış takip sistemleri geliştirenler için çalışmalarına entegre edilebilir yüksek doğrulukta çalışan yöntemlerdir.

Bir diğer çalışmada görüntü işleme teknikleri kullanılarak sadece bir web kamerası ve bir yazılımla çalışan düşük maliyetli bakış takip sistemi önerilmiştir. Bu sistemde bakış takibinin doğruluğunu etkileyen farklı aydınlatma koşullarında gözde oluşan yansımalar, morfolojik işlemler kullanılarak ortadan kaldırılarak bakış takibi gerçekleştirilmiştir.

Günümüz ticari bakış takip cihazları yüksek maliyetinden dolayı geniş bir kullanım alanına sahip değildir. Bu cihazların yüksek maliyetinin yanında farklı ortam koşullarındaki performansı ve uygulama prosedürleri de bu cihazların kullanımını kısıtlayan diğer unsurlardır. Bakış takip teknolojinin herkes tarafından kullanılabilmesi için düşük maliyetli ve web kameralarıyla çalışabilen tüketici sınıfı bakış takip uygulamalarına ihtiyaç vardır. Bu amaçla bir web kamerası ve bir Python yazılımdan oluşan derin öğrenme tabanlı bir bakış takip sistemi geliştirilmiş ve bir web sayfası üzerinde deneyimlenmiştir.

Son olarak, günümüz bakış takip teknolojisinin çocuklarda dikkat eksikliğinin tanısında kullanılabileceğini gösteren, bakış takip ölçüm metrikleri ile eğitilmiş makine öğrenmesi tabanlı bir sistem geliştirilmiştir. Geliştirilen sistem ile klinisyenlerin DEHB olan bireylerin geleneksel tanı süreçlerine yardımcı bir karar destek sistemi önerilmiştir.

1.5 Tezin Organizasyonu ve Yapılan Bilimsel Katkıları

Tez çalışmasının ikinci bölümünde yüksek doğrulukta göz tespitine dayanan alternatif bir yaklaşım sunulmaktadır. Yapılan bilimsel çalışma, “Yüz İmgelerinden Göz Bölgelerinin Tespitinde ESA Tabanlı Alternatif Bir Yaklaşım” başlıklı bilimsel bir makale şekline dönüştürülerek “Fırat Üniversitesi Mühendislik Bilimleri Dergisi” adlı bilimsel dergide yayınlanmıştır.

Üçüncü bölümde bakış takip sistemlerinin önemli bir parametresi olan göz bebeği merkezinin lokalizasyonunu yüksek doğrulukta gerçekleştiren derin öğrenme tabanlı bir yaklaşım sunulmaktadır. Yapılan bilimsel çalışma, “Pupil Center Localization Based on Mini U-Net” başlıklı bilimsel bir makale şekline dönüştürülerek “Journal of Computer Science” adlı bilimsel derginin özel sayısında yayınlanmıştır.

Dördüncü bölümde görüntü işleme tekniklerine dayanan bakış takip sistemi sunulmaktadır. Yapılan bilimsel çalışma, “Video Tabanlı Gerçek Zamanlı Göz Takibi” başlıklı bilimsel bir makale şekline dönüştürülerek “28. IEEE Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU-2020)” adlı konferansta yayınlanmıştır.

Beşinci bölümde yüksek bütçeli ticari bakış takip cihazlarına alternatif sağlam ve yüksek doğrulukta bakış takibi gerçekleştiren bir yaklaşım sunulmaktadır. Yapılan bilimsel çalışma, “A CNN Based Real-Time Eye Tracker For Web Mining Applications” başlıklı bilimsel bir makale şekline dönüştürülerek “Multimedia Tools and Applications” adlı bilimsel dergide yayınlanmıştır.

Altıncı bölümünde bakış takip sistemlerinin çocuklarda dikkat eksikliğinin tespit edilmesine dayanan makine öğrenmesi temelli bir sistem sunulmaktadır. Yedinci bölümde tezin aşamaları genel olarak değerlendirilmiş ve ileride yapılacak çalışmalar için farklı fikirler ele alınmıştır.

2. YÜZ İMGELERİNDEN GÖZ BÖLGELERİNİN TESPİTİNDE ESA TABANLI ALTERNATİF BİR YAKLAŞIM

Tezin bu bölümünde, bakış takip uygulamalarında kullanılabilecek sağlam bir göz tespit yöntemi sunulmuştur. Göz tespiti birçok alanda sistemin giriş verisi olarak kullanılmaktadır. Gözlerin sisteme giriş verisi olarak alındığı alanlara sürücü davranış analizi, bakış takibi, biyometrik güvenlik sistemleri ve göz hastalığı tanıma sistemleri örnek olarak verilebilir [54–59]. Bu sistemlerde gözün algılanabilmesi için öncelikle yüz tespitinin yapılması gerekmektedir. Yüz tespiti yapıldıktan sonra gözün tespiti için farklı yöntemler kullanılabilir. Gözün tespit edilmesinde farklı aydınlatma ortamları, düşük çözünürlüklü görüntüler, kapalı göz, baş pozisyonu gibi faktörler gözlerin algılanmasında önemli bir sorun oluşturur. Bu faktörler göz tespitinin verimli bir şekilde gerçekleştirilmesini zorlaştırmaktadır.

İşlemci hızlarının artmasıyla beraber kullanım alanı yaygınlaşan derin sinir ağları, farklı alanlarda gösterdiği başarılarla performansını ispatlamıştır. Bu bölümde bir imge seçici yardımıyla yüz görüntüleri üzerinden örnekler toplanmıştır. Toplanan bu örnekler, ESA modeline girdi olarak verilmiş ve yüz görüntülerinden göz tespiti gerçekleştirilmiştir. ESA, örüntü tanımadaki etkili bir algoritma olduğundan dolayı tercih edilmiştir. Bu bölümde önerilen ESA modeli ile göz tespitini engelleyen problemlerin model tarafından örnek verilerle öğrenilmesiyle üstesinden gelinmiştir. Önerilen sistemin başarısı günümüzde yüz ve göz tespitinde yaygın bir şekilde uygulanan Viola-Jones algoritmasının XML tabanlı algoritmasıyla karşılaştırılmıştır. Karşılaştırma sonunda önerilen sistemin, baş pozisyonu, kapalı göz gibi gözlerin tespit edilmesinin zor olduğu yüz görüntülerinde daha iyi performans sergilediği görülmüştür.

Bu bölümde öncelikle literatürdeki göz tespiti ile ilgili çalışmalara yer verilmiştir. Daha sonra önerilen ESA tabanlı göz tespit sisteminin eğitimi için gerekli veri seti oluşturulmuştur. Sonrasında göz tespiti için geliştirilen ESA tabanlı sistem, sistemin akış şeması ve model eğitimi anlatılmıştır. Son olarak, önerilen sistemin XML tabanlı yüz ve göz algılama uygulaması ile karşılaştırılması yapılmış ve elde edilen sonuçlar sunulmuştur.

2.1 İlgili Çalışmalar

Literatürde göz tespiti ile ilgili çalışmalar şu şekilde sıralanabilir; Viola ve diğ. [60] yüz ve göz tespitinde etkili Haar-like özellikler, İmge integrali, Basamaklı sınıflandırıcı ve Adaboost bileşenlerinden oluşan XML tabanlı bir nesne tanıma algoritması geliştirmişlerdir. Zhang ve diğ. [61] geliştirdikleri ESA tabanlı Çok Görevli Basamaklı Evrişimli Sinir Ağları (ÇGBESA-MTCNN) yapısı ile yüz ve yüz üzerindeki dönüm noktalarını başarılı bir şekilde tespit etmişlerdir. Geliştirilen yapı P-Net, R-Net ve O-Net aşamalarından oluşmaktadır. Bu aşamalardan P-Net yüz dönüm noktalarının tespiti için ilgili aday pencereleri üretir. Bu pencereler R-Net yardımıyla iyileştirilir. Son olarak O-Net yardımıyla yüzü sınırlayan kutu ve yüz dönüm noktalarının konumu üretilir. Kazemi ve diğ. [62] çalışmalarında yüzün dönüm noktası konumlarını yüksek doğrulukta ve gerçek zamanlı olarak tespit etmek için regresyon ağaçları topluluğunu kullanmışlardır. Eğitiminde HELEN veri kümesini kullandıkları çalışmalarında öğrenim yöntemi olarak da gradyan arttırmayı seçmişlerdir. Soetedjo [63] üç aşamadan oluşan çalışmasında göz tespiti için ilk olarak Viola-Jones algoritmasını kullanarak yüz tespiti gerçekleştirmiştir. Elde edilen yüz bölgesinden kaba bir göz bölgesi elde etmek için yüz bölgesine dikey gradyan uygulamıştır. Dikey gradyan ile elde edilen kaba göz bölgesinden beyaz sklera bölgesini çıkarmak için göz bölgesine renk eşiği uygulamıştır. Son olarak eşik uygulanmış göz bölgesine gözün elips şekline uygun bir elips oturtma yöntemiyle göz sınırlarını tahmin etmeye çalışmıştır. Uygulanan yöntem ile göz sınırı %91'lik bir doğruluk ile tespit edilmiştir. Majumder ve diğ. [64] önerdikleri yöntemde göz konumlarının tespiti için ilk olarak, gri görüntüler üzerindeki aydınlatma efektinin ortadan kaldırılması için kontrast artırma uygulamışlardır. Elde ettikleri görüntülerde göz konumunun tespiti için Hızlı Köşe Dedektörü adı verilen özellik algılama tekniğini kullanmışlardır. Önerdikleri yöntemin doğruluğunu sınamak için rastgele seçilen görüntülerde manuel olarak belirlenmiş göz konumları ile önerdikleri yöntemin tespit ettiği göz konumları karşılaştırılmış ve önerilen yöntemin başarısını ortaya koymuşlardır. Knapik ve diğ. [65] önerdikleri yöntemde araç sürücülerinin yorgunluğunun takip edilmesi amacıyla göz tespitinde farklı aydınlatma koşullarından kaynaklı problemlerin aşılması amacıyla termal görüntüler kullanmışlardır. Önerdikleri yeni sanal yüksek dinamik aralık prosedürü ve bir ön işleme adımıyla termal görüntülerin kontrastı artırılmıştır. Önerdikleri yöntemi ESA tabanlı YOLOv3 nesne dedektörü ile karşılaştırmışlardır. Karşılaştırma sonucunda önerilen yöntemin hız ve doğruluk açısından başarısını ortaya koymuşlardır. Gou ve diğ. [66] sundukları kademeli bir regresyon çerçevesi ile ilk olarak göz algılamayı sonrasında

algılanan gözün durumunu (açık-kapalı) tahmin etmişlerdir. Sentetik ve gerçek görüntülerle eğitilen kademeli regresyon çerçevesi, göz tespiti ve göz durumu tahmini için son teknoloji metotlarla karşılaştırılmış ve çerçevenin performansı ortaya konmuştur. Yu ve diğ. [67] yaptıkları çalışmada gri yoğunluk Varyans Filtresi (VF) ve DVM'nin birleşimine dayalı hibrit bir göz tespit yöntemi önermişlerdir. Önerilen yöntemde görüntülerdeki göz olmayan bölgeler VF yardımıyla görüntüden çıkarılırken DVM yardımıyla da iki göz konumu belirlenmiştir. Ghazali ve diğ. [68] yeni bir göz tespit algoritması geliştirmişlerdir. Eğittikleri göz filtresi ile görüntülerdeki arka plan, ten rengi gibi kısıtlamalardan etkilenmeden gözün konumunu doğru bir şekilde tespit etmişlerdir. Geliştirdikleri algoritma, Gabor filtresi ve OpenCV Cascade sınıflandırıcı ile karşılaştırmış ve diğer sınıflandırıcılara göre daha iyi bir çalışma zamanı performansı elde etmişlerdir. Turan ve diğ. [56] geliştirdikleri bakış takibi uygulamalarında dizüstü bilgisayar kullanıcılarının gerçek dünya koşulları altındaki göz bölgesini Yapay Sinir Ağı (YSA) kullanarak tespit etmeye çalışmışlardır. YSA eğitiminde bilgisayara yerleşik kamera aracılığıyla elde edilen 107 kullanıcıya ait görüntüler kullanılmıştır. Eğitilen YSA modeli ile kullanıcıların gözleri yüksek bir doğrulukla tespit edilmiştir.

2.2 Materyal ve Metodoloji

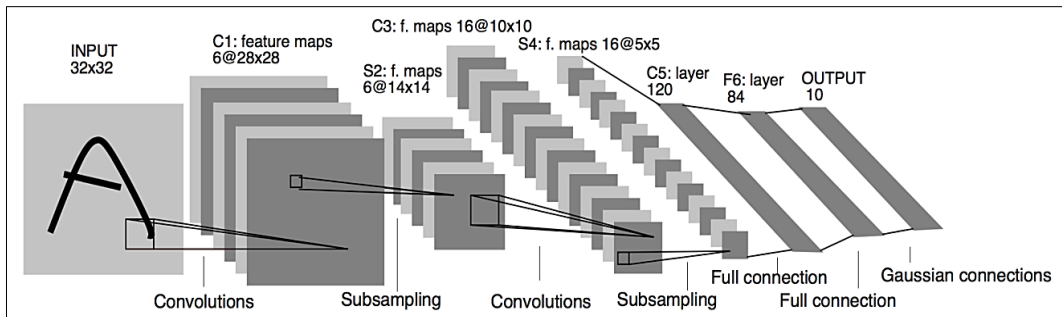
Göz tespitinin derin öğrenme modelleriyle gerçekleştirilebilmesi için derin ağ modellerinin büyük eğitim verileri ile eğitilmesi gerekmektedir. Bu çalışmada eğitim verileri yüksek miktarda yüz görüntülerinden oluşan veri setleri üzerinde gerçekleştirilen görüntü işleme teknikleri ile yüzdeki göz ve göz olmayan bölgeler kırılarak önerilen ESA modeli için veri seti oluşturulmuştur.

2.2.1 Evrimsel Sinir Ağları

1957 yılında Rosenblatt ve arkadaşları tarafından Tek Katmanlı Algılayıcı ile başlayan ve sinir ağlarının tarihsel dönüm noktaları olarak adlandırılabilir yapay sinir ağı serüveni 1960'larda Ivakhnenko ve arkadaşları tarafından çok katmanlı derin ağlarla devam etmiştir [69–71]. 2006 yılında, derin öğrenme terimi, yapay sinir ağlarına yeni bir yaklaşım olan Derin Sinir Ağları (DSA) Hinton ve arkadaşları tarafından tanıtılmıştır [72].

YSA, giriş, gizli ve çıkış katmanlarından ve bu katmanların nöronlarından oluşan bir makine öğrenme yöntemidir [73]. 2000'li yıllarda YSA'daki gizli katman sayısındaki artışa

rağmen yetersiz işlemci hızları nedeniyle yapay zekâ yöntemleri durma noktasına gelmiştir [74,75]. Ancak gelişen teknoloji, yazılım ve donanımdaki gelişmelerle birlikte YSA yöntemleri yeniden ön plana çıkmıştır. YSA'daki katmanların derinleşmesiyle birlikte derin sinir ağları ve derin öğrenme ortaya çıkmıştır. Derin öğrenme, derin sinir ağlarına dayanan makine öğreniminin bir alt dalıdır. Birden fazla eğitilebilir katmana sahiptir ve örüntü tanıma için ham verilerden öznitelikler çıkarabilir [76]. Derin öğrenme, geleneksel YSA yöntemlerinden daha fazla ağ katmanına ve veriye sahip olma eğilimindedir [77]. Literatürde ESA, Uzun Kısa Süreli Bellek (UKSB-LSTM) Ağları, Derin Otomatik Kodlayıcılar, Derin Boltzmann Makineleri (DBM'ler) ve Tekrarlayan Sinir Ağları (TSA) gibi birçok derin öğrenme algoritması bulunmaktadır [78–82]. Derin öğrenmenin gelişmesiyle birlikte ESA, örüntü tanıma problemini çözmede ve makine öğreniminde başarılı bir derin öğrenme algoritması haline gelmiştir. Yapısal olarak ESA'na benzer ilk çalışma 1980 yılında Kunihiro Fukushima tarafından tanıtılmıştır. Fukushima, görüntüler üzerinde çalışan bir sinir ağı olan neocognitron adlı bir model geliştirmiştir [83]. İlk ESA mimarisi, 1988'de LeCun tarafından geliştirilen LeNet ağıdır [78]. 1989 yılında Bell laboratuvarlarında yapılan çalışmada, geriye yayımlı LeNet ağı kullanılarak elle yazılmış sayısal görüntüleri içeren MNIST veri tabanı rakamları doğru bir şekilde sınıflandırılmıştır [84]. LeCun ve diğerleri tarafından geliştirilen LeNet-5 derin öğrenme ağı ile el yazısı rakamlar %99'dan fazla doğrulukla tanınmıştır [85]. LeNet-5 ağı Şekil 2.1'te gösterilmiştir. ESA tabanlı modeller genellikle üç katmandan oluşur; evrişimsel katman, havuz katmanı ve tam bağlantılı katman [85]. İlk evrişim katmanlarında kenarlar ve köşeler gibi daha düşük seviyeli özellikler öğrenilirken, daha derin katmanlarda daha yüksek seviyeli özellikler öğrenilmektedir [86].



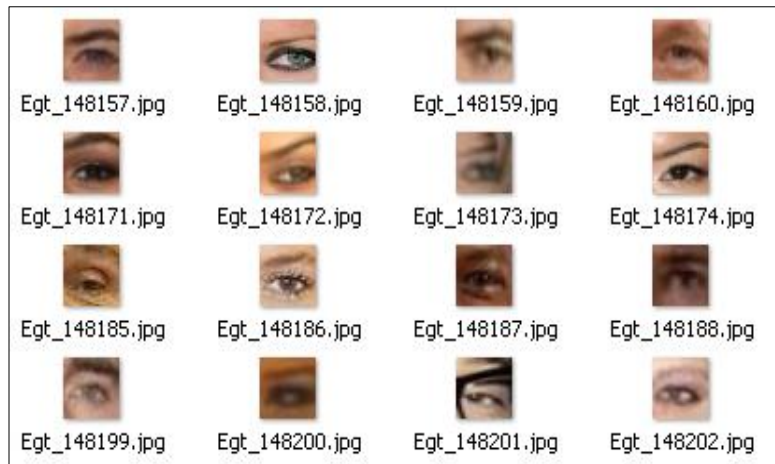
Şekil 2.1: LeNet-5 ESA yapısı [85]

ESA mimarisi ile önemli bir etki yaratan ilk çalışma, 2012 yılında ImageNet ILSVRC-2012 yarışmasında kullanılan ESA tabanlı büyük model olmuştur. Bu model ile

1.2 milyon yüksek çözünürlüklü görüntü 1000 farklı sınıfa ayrılarak rakiplerine karşı yüksek başarı elde etmiştir [87]. ESA'nın örüntü tanımadaki başarısı ile doğal dil işleme, tıbbi görüntü analizi, nesne tanıma, finansal analiz ve birçok farklı alanda ESA tabanlı çalışmalar yapılmıştır [88–91].

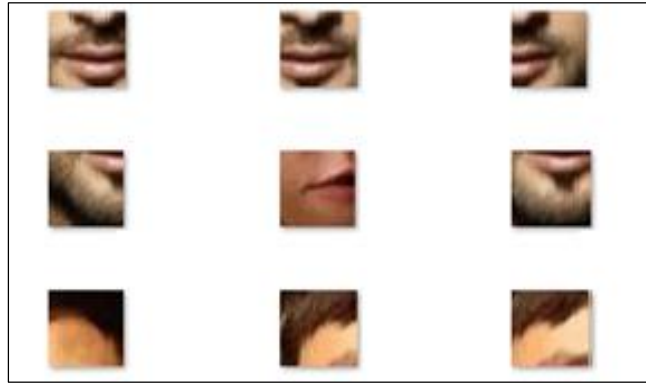
2.2.2 Veri seti oluşturma

Göz tespiti için önerilen ESA modelinin eğitimi için gerekli veriler, 200.000'den fazla tanınmış insan yüzünü barındıran “Large-scale CelebFaces Attributes (CelebA)” adlı veri kümesinden toplanmıştır [92]. Bu veri setindeki görüntülerde farklı baş pozisyonları ve aydınlatma, gözlük, düşük çözünürlük, kapalı göz gibi yüz resimlerinin olması veri setinin seçiminde önemli rol oynamıştır. Veri kümesindeki 100.000 adet yüz görüntüsünden göz bölgeleri, Kazemi ve Sullivan'ın 2014 yılındaki “Regresyon Ağaçları Topluluğu ile Bir Milisaniyelik Yüz Hizalama” makalesinin [62] iBUG 300-W veri kümesi [93] ile eğitilen bir uygulaması olan yüz ve yüz işaret detektörü ile çıkarılmıştır. Bu detektör Dlib kütüphanesinde [94] yer almaktadır. Dlib kütüphanesi, makine öğrenimi ve veri analizi algoritmalarını içeren C++ programlama dilinde yazılmış platform bağımsız, açık kaynak bir araç setidir. Bu işlemden sonra elde edilen 200.000'e yakın 30x30 piksel boyutlarındaki sağ ve sol göz görüntülerinden sınırlı işlemci kaynakları göz önünde bulundurularak 150.000 adet veri kullanılmıştır. Elde edilen görüntüler içerisinde gözlüklü, gözlüksüz görüntüler mevcuttur. Ayrıca, bu görüntüler farklı açılardan alınmış olup çeşitli parlaklık ve çözünürlükler içermektedir. Yüz görüntülerinden kırılmış göz görüntüleri Şekil 2.2'de verilmiştir.

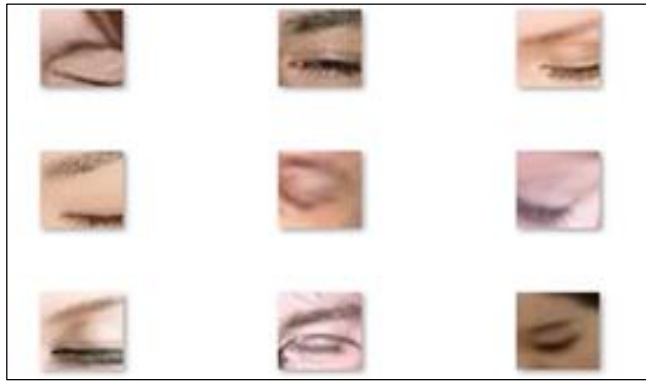


Şekil 2.2: Kırılmış göz görüntüleri

CelebA veri kümesinden çıkarılan göz bölgelerinin önerilen ESA modeli tarafından öğrenilebilmesi için yüzde göz olmayan bölgelere de ihtiyaç duyulmaktadır. Bu sebeple aynı veri setinden 30x30 piksel boyutlarında 100 bin adet yüz bölgesinde gözü içermeyen bölgeler alınmıştır. Yüz görüntülerinden göz olmayan kırılmış bölgeler Şekil 2.3'te verilmiştir. Son olarak kapalı gözlerin de ESA modeli tarafından doğru bir şekilde öğrenilebilmesi için 1192 görüntüden oluşan "Closed Eyes In The Wild (CEW)" adlı veri kümesinden 30x30 piksel boyutlarında sağ ve sol olmak üzere 1600 adet kapalı göz görüntüsü toplanmıştır [95]. Veri kümesindeki yüzlerin bir bölümü yüz dedektörü tarafından algılanamadığından bu yüzlere ait kapalı gözler eğitim verilerine dahil edilememiştir. Şekil 2.4'te kapalı gözlere ait görsel verilmiştir.



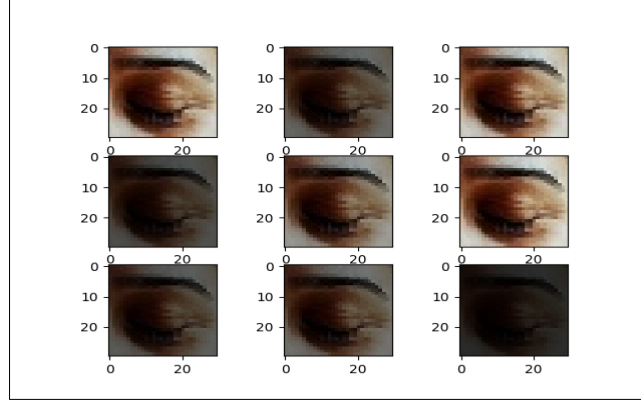
Şekil 2.3: Yüzde göz olmayan bölgeler



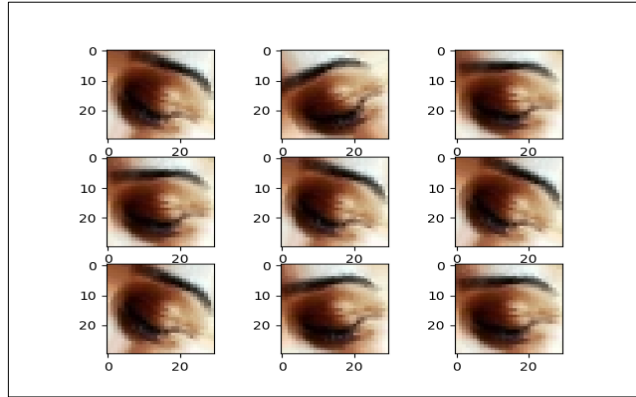
Şekil 2.4: Kapalı göz görüntüleri

Modelin eğitim aşamasında kapalı gözlerin model tarafından daha iyi öğrenilmesi için mevcut 1600 adet kapalı göz görüntüsü yeterli olmamıştır. Bu yüzden 1600 kapalı göz görüntüsü üzerine veri arttırma yöntemlerinden 20°'lik rotasyon ile rastgele parlaklık uygulanarak kapalı göz veri sayısı 10.000'e çıkarılmıştır. Kapalı gözlere ait veri setinde

gerçekleştirilen veri artırımına yönelik görsel Şekil 2.5 ve Şekil 2.6’da gösterilmiştir. Kapalı gözlere ait verilerle beraber toplamda 260 bin adet veri üzerinde çalışılmıştır. Bu verilerden açık-kapalı göz görüntüleri için “1”, göz içermeyen görüntüler için “0” etiketi kullanılmış ve verilerin büyük bir bölümü ESA modeli ile eğitilmiştir.



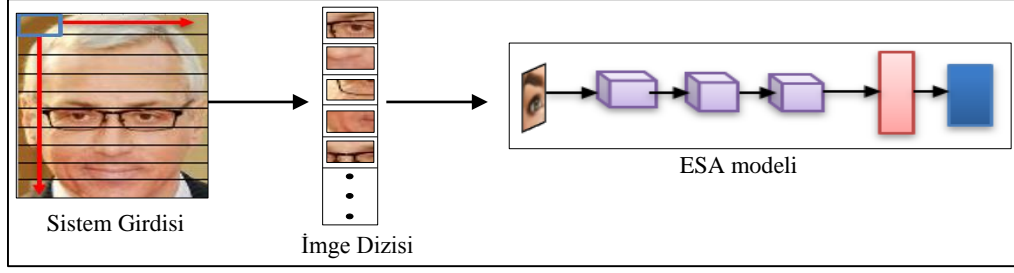
Şekil 2.5: Rastgele parlaklık



Şekil 2.6: Rastgele rotasyon

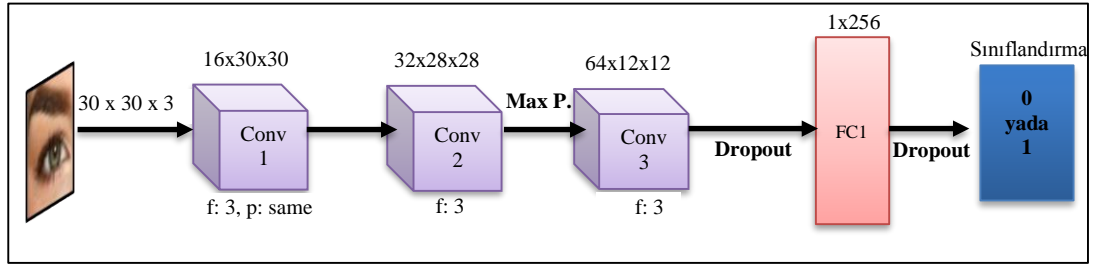
2.2.3 Önerilen ESA tabanlı sistem ve akış şeması

Önerilen sistemde ESA tabanlı model tarafından göz tespitin gerçekleştirilebilmesi için girdi olarak verilen 100x100 uzamsal boyutlarındaki yüz görüntüsü kullanılmıştır. Bu görüntüler üzerinden 30x30 uzamsal boyutta örnekler çıkarılmıştır. Örneklerin oluşturulmasında imge seçiciden yararlanılmıştır. İmge seçici yardımıyla toplanan 30x30’luk imgeler bir dizi değişkeninde biriktirilerek sınıflandırılmak üzere eğitilen ESA modeline verilmiştir. ESA modeli bu imgeleri “göz” veya “göz değil” olarak sınıflandırmaktadır. Önerilen sisteme ait görsel Şekil 2.7’de verilmiştir.



Şekil 2.7: Önerilen ESA tabanlı sistem

Önerilen sisteme ait ESA modelinde $30 \times 30 \times 3$ uzamsal boyutlara sahip göz görüntüleri 2 evrişim katmanından geçtikten sonra elde edilen çıktıya maksimum havuzlama (max pooling) uygulanmıştır. Havuzlama katmanından sonra tekrar evrişim uygulanmıştır. Evrişim katmanının hemen ardından dropout uygulanmıştır. Dropout katmanından sonra elde edilen 3 boyutlu matris düzleştirilerek tam bağlantı katmanına verilmiştir. Tam bağlantı katmanından sonra elde edilen çıktıya yine bir dropout uygulanarak sınıflandırma katmanına verilmiştir. Katmanların aktive edilmesi için tüm katmanlarda ReLU, son katmana ise ikili sınıflandırma için softmax aktivasyon fonksiyonu uygulanmıştır. Şekil 2.8’de Sistem ESA modeli gösterilmiştir.

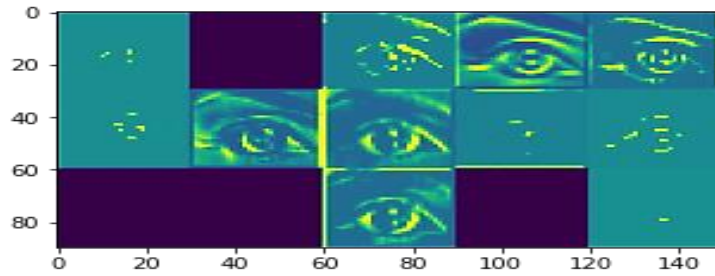


Şekil 2.8: Önerilen sistemin ESA modeli

Şekil 2.8’deki 1, 2 ve 3 nolu evrişim katmanları $30 \times 30 \times 3$ uzamsal boyutlara sahip göz giriş görüntüsüne sırasıyla uygulanmıştır. $3 \times 3 \times 3$ uzamsal boyutta ve 0, 1, -1’den oluşan filtreler, görüntü üzerinde belirli aralıklarla kaydırılarak evrişim işlemi gerçekleştirilmiştir. Bu işlemler sonunda, uygulanan filtrenin boyutuna bağlı olarak çıktının uzamsal boyutları değişmiştir. Evrişim katmanından sonra elde edilen çıkışlar bir sonraki evrişim katmanı için giriş verisi olarak hazırlanmaktadır. Şekil 2.8’de görüntülere uygulanan evrişim işleminin hesaplanması Denklem 2.1’de verilmiştir [96].

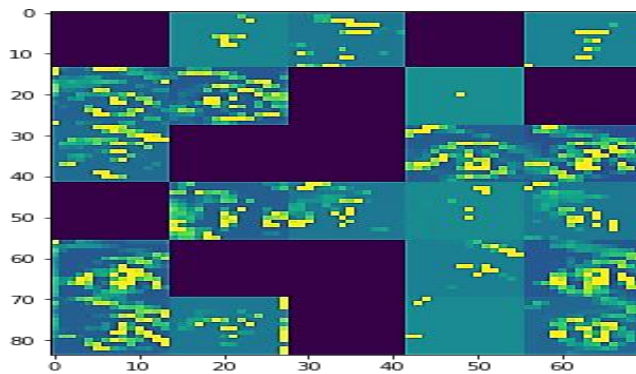
$$F_c = \text{conv}(i, j) = (I \otimes K)(i, j) = \sum_m \sum_n I(m, n)k(i - m, j - n) \quad (2.1)$$

Denklem 2.1'deki F_c , \otimes , m ve n ifadeleri sırasıyla evrişim çıktısına, evrişim işlemine ve görüntünün uzamsal boyutlarına karşılık gelmektedir. Denklemde verilen I girdi görüntüsüne, K filtre (çekirdek)'ye karşılık gelmektedir. K filtresinin I üzerinde soldan sağa ve yukarıdan aşağıya doğru kaydırılmasıyla girdi görüntü matrisinin her bir elemanı çekirdeğin her bir elemanı ile çarpılır. Görüntü matrisi üç katmanlı ise bu işlem her bir katman için ayrı ayrı gerçekleştirilir ve elde edilen çıktılar evrişim çıktısının oluşturulması için birleştirilirler [96]. Sistem ESA modelinin ilk evrişim katmanı çıktısı Şekil 2.9'da gösterilmiştir.



Şekil 2.9: 1 nolu evrişim katmanı çıktıları

ESA modelindeki hiper parametrelerden maksimum havuzlama aşamasında giriş verilerine filtre (2x2) uygulanmış ve giriş verilerinin maksimum değerleri seçilerek görüntü özelliklerine ait uzamsal bir alt boyut elde edilmiştir. Bu işlem hesaplama kolaylığı sağladığı gibi önemli özelliklerin çıkarılmasında da rol oynar [97]. Modelin havuzlama katmanı çıktıları Şekil 2.10'da gösterilmiştir. Görüntü özelliklerine uygulanan bir sonraki parametre olan Atma (Dropout), katmanlarda yer alan nöronların belirlenen oranda modelden rastgele çıkarılmasıyla gerçekleşir. Atma işlemi modelin eğitimi sırasında, aşırı uydurma (overfitting)'nin üstesinden gelmek ve yeni veriler için modelin genelleme gücünü arttırmak için kullanılmaktadır [98,99].



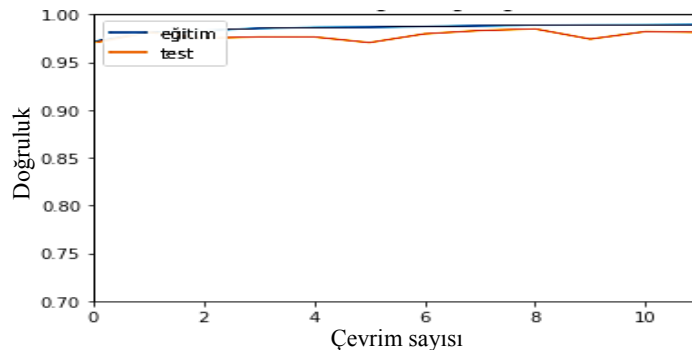
Şekil 2.10: Maksimum havuzlama katmanı çıktısı

2.2.4 Sistem parametreleri ve eğitim

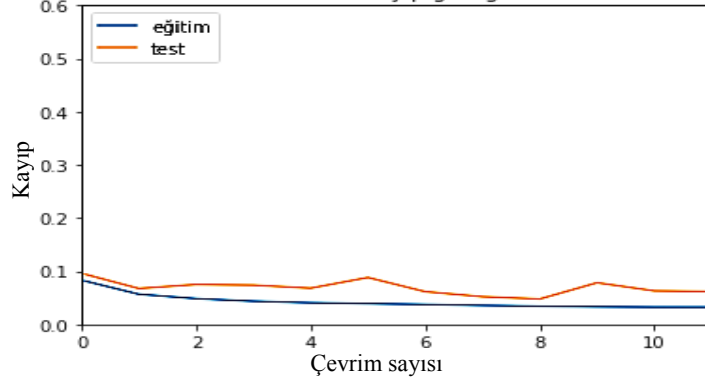
Açık-kapalı göz toplamda 160.000 adet ve yüzde göz olmayan imgeler 100.000 adet olmak üzere toplamda 260.000 veri üzerinde modelin eğitim ve test işlemleri gerçekleştirilmiştir. 260.000 verinin %80'i eğitim aşamasında %20'si test aşamasında kullanılmıştır. Eğitim aşamasında doğruluk %99 iken test aşamasında bu oran %97,97 ile sonuçlanmıştır. Elde edilen sonuçlar, ağın bu verileri doğru bir şekilde öğrendiğini göstermektedir. ESA modeli verileri Çizelge 2.1'de verilmiştir. Eğitim aşaması 15 çevrimde gerçekleşmiştir. Her çevrimde veri seti 32 parti boyutu halinde eğitilmiştir. ESA modelinin optimizasyonu “Adam (lr=0.0001)” optimizasyon algoritması, kayıp fonksiyonu olarak da “categorical crossentropy” kullanılmıştır. Şekil 2.11 ve 2.12'de sırasıyla ESA modelinin doğruluk ve kayıp grafikleri verilmiştir.

Çizelge 2.1: Sistem ESA modeli verileri

Sistem ESA modeli verileri	
Veri seti	260.000
Eğitim verisi	% 80
Test verisi	% 20
Giriş verileri	Açık-Kapalı göz imgeleri (150.000 + 10.000) Yüzde göz olmayan imgeler (100.000)
Kayıp fonksiyonu	Categorical crossentropy
Optimizasyon	Adam
Parti boyutu	32
Eğitim döngü sayısı	15
Doğruluk (Eğitim)%	99.00
Doğruluk (Test)%	97.97



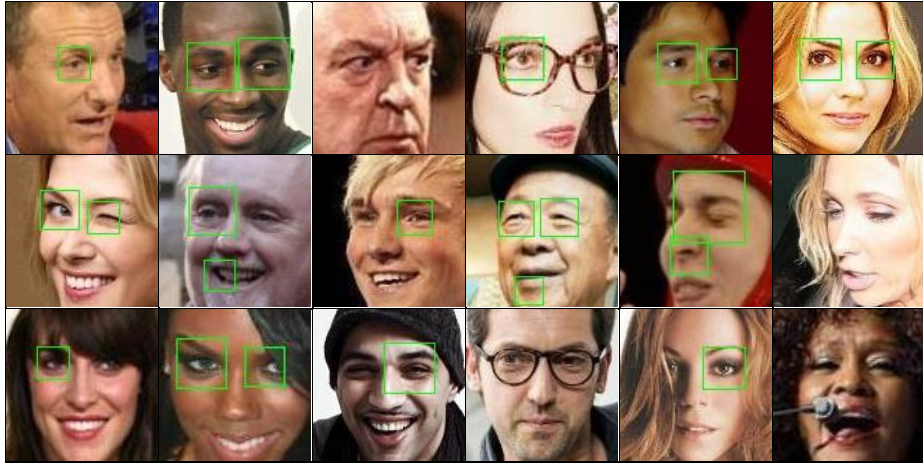
Şekil 2.11: Model doğruluk grafiği



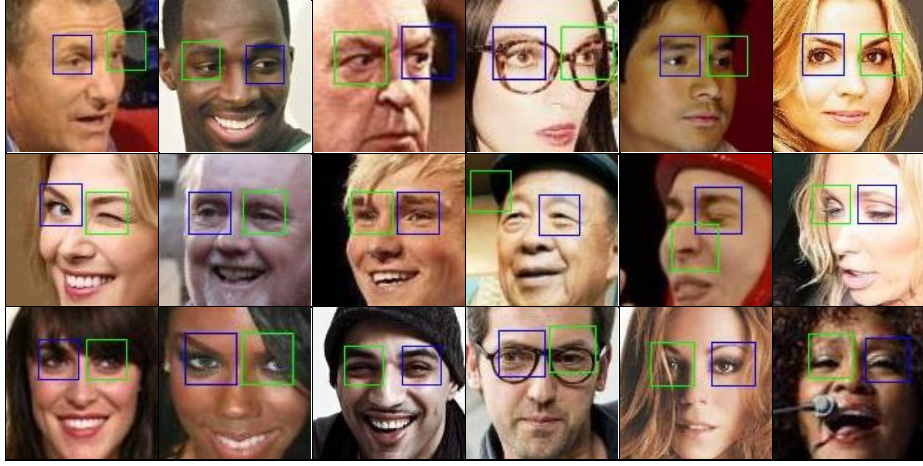
Şekil 2.12: Model kayıp grafiği

2.3 Deneysel Sonuçlar

Önerilen sistemin performansının ölçülmesi için sisteme verilen test görüntüleri üzerinde imge seçici yardımıyla elde edilen imgeler ESA modeline verilir. Evrişim işlemlerinden geçen imgeler sınıflandırma katmanında softmax aktivasyon fonksiyonu ile aktive edilerek 0-1 arasında tahmini değerlere sahip olurlar. Bu imgelere ait tahmin değeri 1'e yakın ise "göz" (kapalı-açık), 0'a yakın ise "göz değil" olarak kabul edilmiştir. Yüz görüntüsü üzerinde göz olarak tahmin edilen imgeler ikiden fazla olabilir. Bunu sebebi olarak yüz üzerinde imge toplama işlemi esnasında imge seçici için uygulanan kayma miktarına bağlı olarak göz bölgesinden birden fazla örnek alınabilmektedir. Bu sorunun üstesinden gelmek için tahmin sonunda 1 değerine en yakın skora sahip imgeler göz olarak kabul edilmektedir. Böylelikle her iki göz için en yüksek skor göz imgesi olarak kabul edilir. Şekil 2.13 ve Şekil 2.14'te CeleBA veri setinin, eğitimde kullanılmayan yüz görüntülerinin sırasıyla Viola Jones algoritması ve önerilen ESA tabanlı sistem üzerinde uygulanması ile elde edilen sonuçların bir kısmı gösterilmiştir.



Şekil 2.13: Viola-Jones algoritması uygulanan görüntüler



Şekil 2.14: Önerilen ESA tabanlı sistemin uygulandığı görüntüler

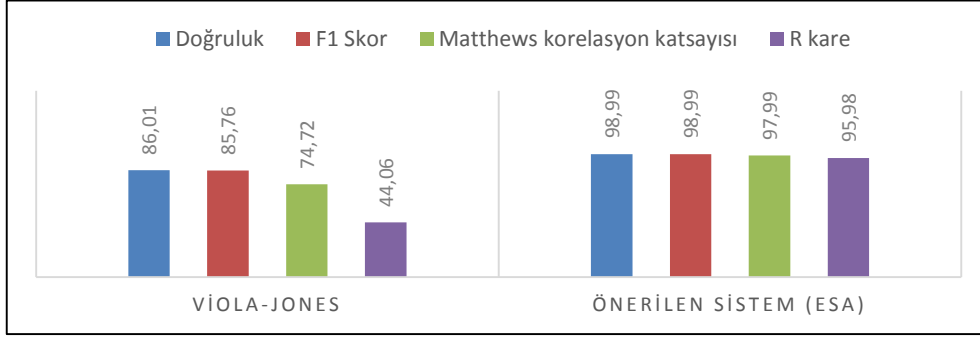
Karşılaştırma için veri setinden görüntüler seçilirken sağa, sola, öne bakan ve farklı çözünürlüklere sahip yüz görüntülerinin olmasına özellikle dikkat edilmiştir. Ayrıca bu resimlerde kapalı göz, gözlüklü, saç ile gizlenmiş göz durumlarının kullanılmasına sistemin performansı açısından ayrıca dikkat edilmiştir. Viola Jones algoritmasının uygulanmasında OpenCV kütüphanesi kullanılmıştır. Kütüphanenin detectMultiScale fonksiyonundaki göz tespit parametrelerinden scaleFactor parametresinin değeri 1.2 ve minNeighbors parametresinin değeri 3 olarak belirlenmiştir [100]. Şekil 2.13 ve 2.14 karşılaştırıldığında önerilen ESA tabanlı sistemin hemen hemen tüm karşılaştırma görüntüsünde iki gözün tespiti doğru bir şekilde gerçekleştirilmiştir.

Önerilen sistemin performansını göstermek için ayrıca bir karşılaştırma yapılmıştır. Eğitim aşamasında kullanılmayan 600 görüntü seçilmiştir. Seçilen görüntülerde gözlüklü yüzlerin yanında farklı duruş pozisyonlarına sahip yüz görüntülerinin seçilmesine dikkat edilmiştir. Seçilen 600 görüntü üzerinde yapılan karşılaştırma sonuçları Çizelge 2.2’de gösterilmiştir. Çizelge 2.2’de, önerilen sistemin 600 görüntünün tamamı üzerinde en az bir göz tespiti yaptığı görülmüştür. Önerilen sistemde iki göz tespitinde %98,2’lik bir başarı sağlanırken Viola-Jones algoritmasında bu oran %84,5’tir.

Çizelge 2.2: 600 görüntü üzerinde karşılaştırma

Viola-Jones ve Önerilen sistem karşılaştırması				
Yöntem	Tek Göz	İki Göz	Hiçbiri	Toplam resim
Viola-Jones	61	507	32	600
Önerilen yöntem	11	589	0	600

Önerilen sistem, doğruluk ölçütleriyle yapılan testlerden de başarıyla geçmiştir. Bu testlerde daha önce kullanılmamış 20.000 görüntü kullanılmıştır. Bu görüntülerin 10.000'i göz görüntüsü diğer 10.000'i ise yüzde göz olmayan görüntülerdir. Bu görüntüler Doğruluk, F1 skor, Matthews korelasyon ve R kare ölçütleriyle sınanmıştır. Performans ölçütlerinin uygulanması python sklearn kütüphanesi fonksiyonları ile gerçekleştirilmiştir [101]. Ölçütlere ait sonuçlar Şekil 2.15'te gösterilmiştir.



Şekil 2.15: Viola-Jones ve Önerilen sistem doğruluk karşılaştırması

Önerilen yöntemde sistemin 30x30 boyutlarında göz araması yapması için yüz görüntülerine boyut sınırlaması getirilmiştir. Bu sınırlama Viola-Jones algoritmasına göre dezavantajlı görünmektedir. Bunun sebebi Viola-Jones algoritması ile farklı boyutlardaki yüz görüntülerinden de gözlerin tespit edilebilmesidir. Önerilen yöntemin farklı boyutlardaki yüz görüntülerinde de göz tespitini gerçekleştirebilmesi için yüz görüntülerinin uzamsal boyutlarının görüntü işleme teknikleri aracılığıyla artırılması veya azaltılması sonucunda görüntünün 30x30'luk bir çerçeveye göz arama tespitine uygun duruma getirilmesi uygun görülmüştür.

Tezin bu bölümde yapılan çalışmada ESA'nın görüntü tanımadaki gücünden yararlanılarak bakış takip sistemleri için bir göz tespit sistemi önerilmiştir. Önerilen sistemin göz algılamadaki başarısı Viola-Jones algoritması ile karşılaştırılmıştır. Deneysel sonuçlardan elde edilen verilere göre önerilen sistemin göz tespitinde daha iyi performans gösterdiği görülmektedir. Gözlerin tespit edilmesinde çalışma zamanı performansı olarak Viola-Jones algoritmasına kıyasla önerilen sistemin çalışma zamanı performansı, görüntü üzerinde göz adayları imgelerin toplanması sürecinde kaybedilen süreden dolayı daha düşüktür. Bu problemin çözümü olarak, görüntü üzerindeki imgeler toplanırken gerçekleşen kayma miktarının artırılması sayesinde kaybedilen sürenin düşürülmesi ile mümkündür. Sürenin düşürülmesine karşılık sistemin performansı da belirli oranda düşmektedir. Bu durumda yüzün spesifik özelliklerinden yararlanılarak yüzün tamamında gerçekleşen imge

toplama işlemi yerine gözlerin bulunduğu konumlarda imge toplama işleminin gerçekleştirilmesiyle önerilen yöntemin hızı büyük ölçüde artırılabilir. Ayrıca sistemin mevcut haliyle gerçek zamanlı uygulamalarda kullanılması için daha hızlı işlemciler kullanılması gerekmektedir.

3. MİNİ U-NET MİMARİSİNE DAYALI GÖZ BEBEĞİ MERKEZİ LOKALİZASYONU

Tezin bu bölümünde, bakış takibinde ikinci aşama olan göz bebeği merkez lokalizasyonunu için U-Net mimarisine dayalı bir ağ mimarisi önerilmiştir. Göz bebeği merkezinin lokalizasyon tespiti, gözün bakış yönünü tespit etmede etkin olduğu için hala bakış takip sistemlerinde popüler bir araştırma konusudur. Göz bebeğinin/merkezinin doğru bir şekilde tespit edilmesi bakış takip teknolojisinin en önemli aşamalarından biridir. Günümüzde video tabanlı ticari bakış takip cihazları kızılötesi ışık ve yüksek çözünürlüklü kameralar ile yüksek doğrulukta göz bebeği merkezi lokalizasyonu gerçekleştirmektedir. Kullandıkları kızıl ötesi ışık ile gözün iris, göz bebeği ve sklera kısımları net bir şekilde ayırt edilebilmektedir. Dolayısıyla göz bebeği merkezi daha sağlam bir şekilde tespit edilebilmektedir. Fakat bu cihazlar farklı aydınlatma koşullarına karşı çok hassas değildir. Karanlık ortamlarda, yüksek ışık ortamlarına göre daha iyi lokalizasyon tespiti gerçekleştirmektedirler. Bu cihazların yüksek bütçeli olmaları ve bakış takibinin kontrollü koşullarda altında gerçekleştirilmesi kullanıcı deneyimini sınırlamaktadır. Bu sınırlamaların üstesinden gelmek için zorlu gerçek dünya koşullarında, düşük donanımlarda ve düşük çözünürlüklü kameralardan bakış takibi için sağlam bir göz bebeği lokalizasyonu/merkezi elde edilmesi ihtiyacı doğmuştur.

Bakış takip sistemlerinin yüksek doğrulukta çalışmasında etkili olan sağlam bir göz bebeği merkezi lokalizasyonu için geçmişten günümüze birçok yöntem kullanılmıştır. Bu yöntemler genel hatlarıyla ikiye ayrılabilir. Bunlar şekil ve görünüm tabanlı yöntemlerdir. Şekil tabanlı yöntemlerde iris ve göz bebeği lokalizasyonu geleneksel morfolojik görüntü işleme teknikleri ve gözün geometrik özellikleri kullanılarak yapılırken görünüm tabanlı yöntemlerde ise gözü temsil eden spesifik özellik vektörleri çıkarılarak makine öğrenmesi ile iris ve göz bebeği lokalizasyonu gerçekleştirilmektedir. Geleneksel teknikler ışık, gözlük, iris rengi, düşük çözünürlük gibi doğal koşullarından olumsuz etkilenirken görünüm tabanlı yöntemler olumsuz koşulları da birer özellik olarak algılamasından dolayı bu durumdan daha az etkilenmektedir. Bu çalışmada göz bebeği lokalizasyonu için göz özelliklerinin otomatik olarak öğrenildiği ve göz bebeği merkezi lokalizasyonun yüksek doğrulukta yapıldığı

görünüm tabanlı bir yöntem önerilmiştir. Bu yöntem U-Net mimarisine dayalı Mini U-Net ağıdır. Önerilen ağ, göz bebeği merkezi lokalizasyonu için halka açık GI4E veri seti üzerinde değerlendirilmiştir. Ağın test sonuçlarında maksimum normalleştirilmiş hata kriterine göre göz bebeği merkezi %98.40 doğruluk oranıyla tespit edilmiştir. Elde edilen sonuç, güncel teknolojik yöntemlerle karşılaştırılmıştır. Karşılaştırma sonuçları, önerilen ağ mimarisinin sağlam ve doğru bir yöntem olduğunu kanıtlar niteliktedir.

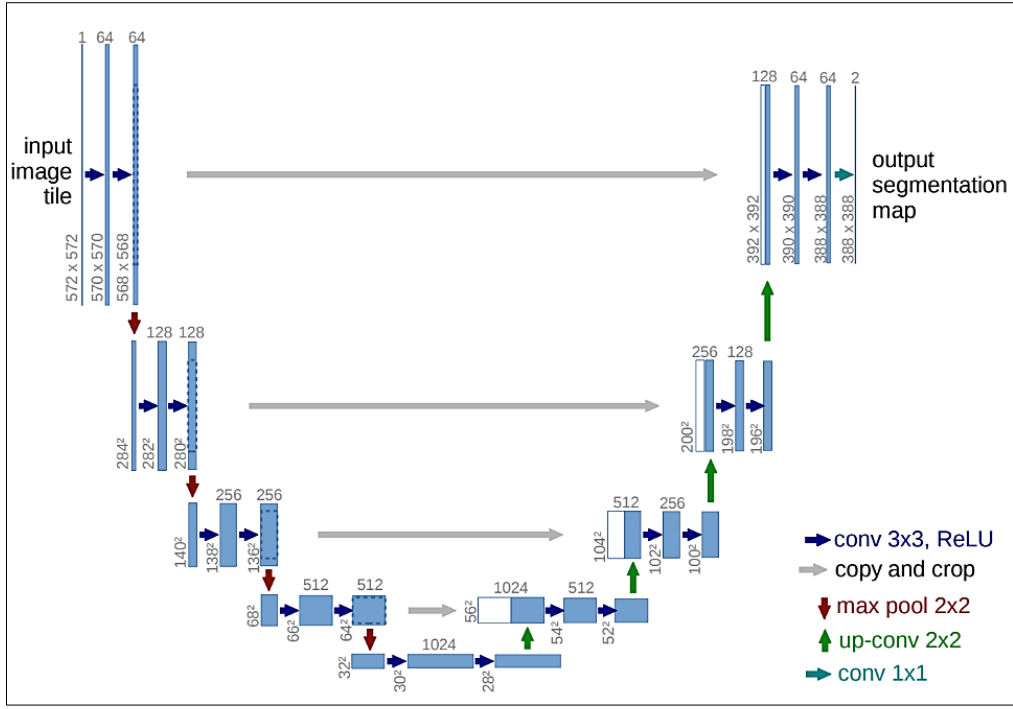
3.1. İlgili Çalışmalar

Gerçek dünya koşullarında göz bebeği lokalizasyonu için birçok algoritma geliştirilmiştir. Cai ve diğ. [102] yaptıkları çalışmada göz merkezi lokalizasyonunda karşılaşılan zorlukların üstesinden gelmek için hiyerarşik uyarlamalı bir evrişim yöntemi önermişlerdir. Bu yöntemdeki yeni hiyerarşik çekirdekler ile irisin farklı görüş açıları modellenmiştir. Hiyerarşik çekirdekler 3B baş pozuna göre uyarlanabilir şekilde seçilmiştir. Böylelikle göz merkezi yerelleştirmesi doğru ve hızlı bir şekilde yapılmıştır. Önerilen yöntemin performansı en sık kullanılan göz yerelleştirme veri kümesi olan GI4E veri setinde denenmiştir. Normalleştirilmiş hata kriterlerinde ($e \leq 0.025$, $e \leq 0.05$) sırasıyla %85.7 ve %99.5 doğruluk elde edilmiştir. Kitazumi ve Nakazawa [103] koyu renk irislerde ve kornea yansıma senaryolarında uygulanabilen ESA tabanlı U-Net mimarisini kullanarak göz bebeği segmentasyonunu gerçekleştirmişlerdir. Önerilen mimarinin eğitiminde renk ve kornea yansıması gibi veri arttırımının yapıldığı UBIRIS.v2 göz veri kümesi ve mobil ortamlarda elde ettikleri göz görüntüleri kullanılmıştır. Eğitilen U-Net mimarisinin göz bebeği merkezi tespiti performansı için GI4E yüz görüntü seti kullanılarak göz bebeği merkezi algılama doğruluğu test edilmiştir. Test sonucunda normalleştirilmiş hata kriterlerinde ($e \leq 0.025$, $e \leq 0.05$, $e \leq 0.1$) sırasıyla %96.28, %98.62 ve %98.95 doğruluk elde edilmiştir. Choi ve diğ. [104] yaptıkları çalışmada göz bebeği lokalizasyonu için heterojen ESA modelleri kullanan bir algoritma önermişlerdir. Önerilen algoritmanın eğitiminde nicel değerlendirme için 5 katlı çapraz doğrulama yöntemini kullandıkları BioID ve GI4E ver kümeleri kullanılmıştır. Önerilen algoritma iki adımdan oluşmaktadır. Bu adımlar göz bölgesi çıkarma ve göz bebeği merkezi lokalizasyonudur. Göz bölgesi tespit edildikten sonra göz bebeği segmentasyonu için Tam Evrişimli Ağ (TEA-FCN) kullanılmıştır. Segmentasyon sonunda elde edilen piksel bazındaki sınıflandırmada maksimum piksel yoğunluğuna sahip piksel konumu göz bebeği merkezi olarak kabul edilmiştir. Ağın doğruluğunun arttırılması için ise TEA ağında kaybolan konum bilgisinin telafi edilmesi için bir atlama bağlantısı ve otomatik kodlayıcının

kod çözücüsüne karşılık gelen yardımcı ağ kullanılmıştır. Bu yardımcı ağ ve TEA ağının kayıpları birleştirilerek doğruluk arttırılmaya çalışılmıştır. Önerilen ağ BioID ve GI4E veri kümesinde test edilmiştir. GI4E veri seti ile yapılan testte göz bebeği merkezi, normalleştirilmiş hata kriterlerinde ($e \leq 0.025$, $e \leq 0.05$) sırasıyla %90.40 ve %99.60 doğruluk oranlarında tespit edilmiştir. Larumbe-Bergera ve diğ. [105] göz merkezi lokalizasyonu için ImageNet veri seti ile önceden eğitilmiş ResNet-50 ağını kullanmışlardır. Eğitiminde PUPPIE veri setinin kullanıldığı ve veri arttırmanın uygulandığı ResNet-50 ağına göz merkezi lokalizasyonuna uygun ince ayar prosedürü uygulamışlardır. Ağın çıkışında iki göze ait göz köşeleri ve merkezleri olmak üzere toplamda 6 dönüm noktası tahmin etmişlerdir. Önerilen ağın performansı GI4E, I2Head, MPIIGaze, U2Eyes veri setlerinde test edilmiştir. Test sonucunda normalleştirilmiş hata kriterlerinde ($e \leq 0.025$, $e \leq 0.05$, $e \leq 0.1$) sırasıyla %98.46, %100 ve %100 doğruluklar elde etmişlerdir.

3.2 U-Net Mimarisi

U-Net mimarisi 2015 yılında Ronneberger ve diğ. [106] tarafından biyomedikal görüntü segmentasyonu için geliştirilmiştir. Bu mimari “U” şeklinde bir yapıya sahiptir. Mimarinin ilk kısmı kodlayıcı ikinci kısmı ise kod çözücü bloklardan meydana gelmektedir. İlk kısımdaki 4 bloğun her bloğunda evrişim işlemleri (3x3) ile görüntünün kanal sayısı iki katına çıkarılarak özellik haritaları elde edilir. Her evrişim aşamasında ReLU aktivasyon fonksiyonu ile tüm nöronlar aktive edilerek ağı doğrusal olmayan durumlara hazır hale getirerek ağın genelleştirme yeteneği arttırılır. Evrişimler sonunda görüntünün uzamsal boyutları maksimum havuzlama (2x2) ile yarıya düşürülmesiyle eğitilebilir parametreler azaltılarak maliyet düşürülür. Kod çözücü blokta ise bu işlem simetrik olarak tersine işlemekte ve 4 bloğun her bloğunda ilk olarak üst örnekleme ile uzamsal boyutlar iki katına çıkarılırken kanal sayısı yarıya düşürülür. Üst örnekleme ters evrişiminden (2x2) sonra kodlayıcı bloğundan atlama bağlantısı ile gelen tensör üst örneklenen tensör ile birleştirilir. Buradaki amaç ağın derinleşmesi ile kaybolan özellikleri telafi etmek ve önceki katmanlardan gelen bilgilerle daha iyi bir tahmin yapabilmektedir. Birleştirme işleminden sonra evrişim işlemleri (3x3) ile kanal sayısı yarıya düşürülmektedir. Her evrişimden sonra ReLU aktivasyon fonksiyonu ile nöronlar aktive edilir. Kod çözücünün son evrişim bloğundaki son evrişim katmanında 1x1 bir evrişim ve son olarak sigmoid aktivasyonu ile piksel düzeyinde sınıflandırmayla segmentasyon gerçekleştirilir. U-Net mimarisi Şekil 3.1’de gösterilmiştir.



Şekil 3.1: U-Net mimarisi [106]

3.3 Materyal ve Metodoloji

Çalışmada göz merkezi lokalizasyonun da yaygın olarak kullanılan GI4E veri seti kullanılmıştır. Veri seti bir dizi ön işlemde geçirilerek önerilen Mini U-Net mimarisinin eğitimine hazır hale getirilmiştir.

3.3.1 Veri seti

Çalışmada göz bebeği bölütlemesinde yaygın olarak kullanılan GI4E veri seti kullanılmıştır. GI4E veri seti, standart web kamerası ile 103 farklı kullanıcının ekranda farklı noktalara baktığı 12 görüntüye karşılık gelen toplamda 1236 yüz görüntüsünden oluşan bir veri setidir. Görüntüler 800x600 piksel çözünürlüğüne sahip png formatındadır. Ayrıca veri tabanı yüz görüntülerindeki iris merkezi ve göz köşe noktalarını her görüntü için toplamda 6 koordinat çifti ile etiketleyen bir metin dosyasına sahiptir. Etiketleme işlemi manuel olarak gerçekleştirilmiştir. Bu işlem üç farklı kullanıcı tarafından iris merkezi ve göz köşelerinin her görüntü üzerinde işaretlenmesi ve bu üç kullanıcının işaretlerinin ortalamaları alınarak gerçekleştirilmiştir [107]. GI4E veri setine ait görüntülerin bir kısmı ve yine aynı veri setinde yer alan gözdeki yer işaretlerinin gösterimi sırasıyla Şekil 3.2 ve Şekil 3.3'te gösterilmiştir.



Şekil 3.2: GI4E veri seti görüntüleri

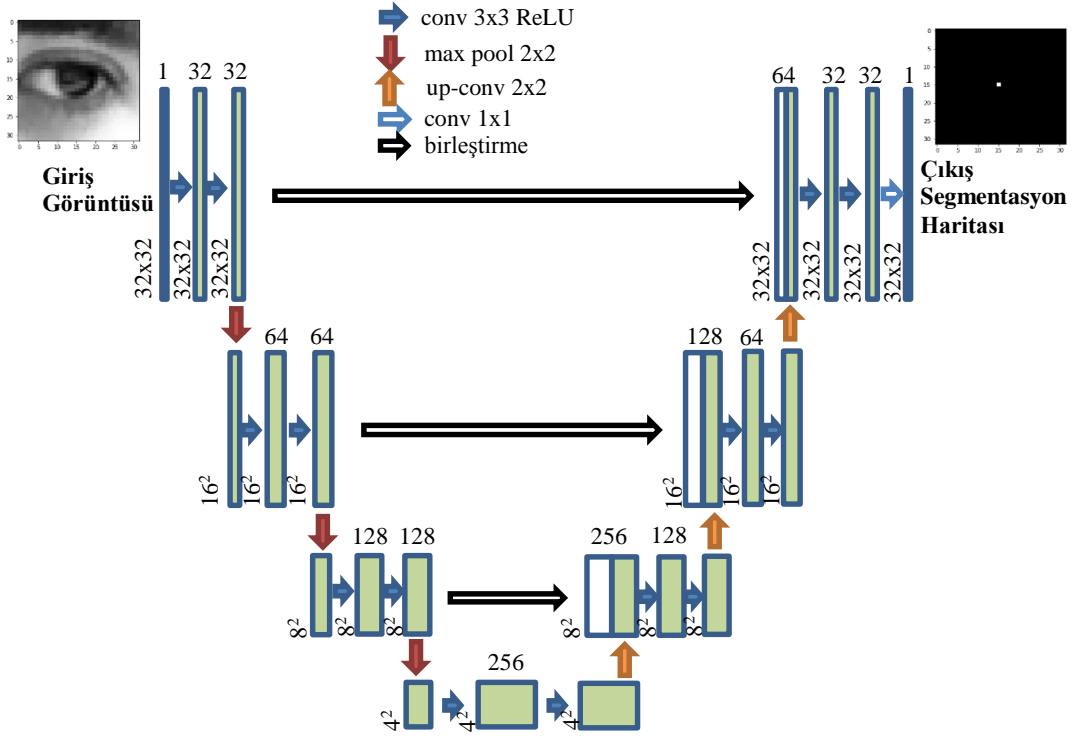


Şekil 3.3: GI4E veri seti görüntüsüne yer işaretlerinin uygulanması

3.3.2 Önerilen Mini U-Net mimarisi ve eğitimi

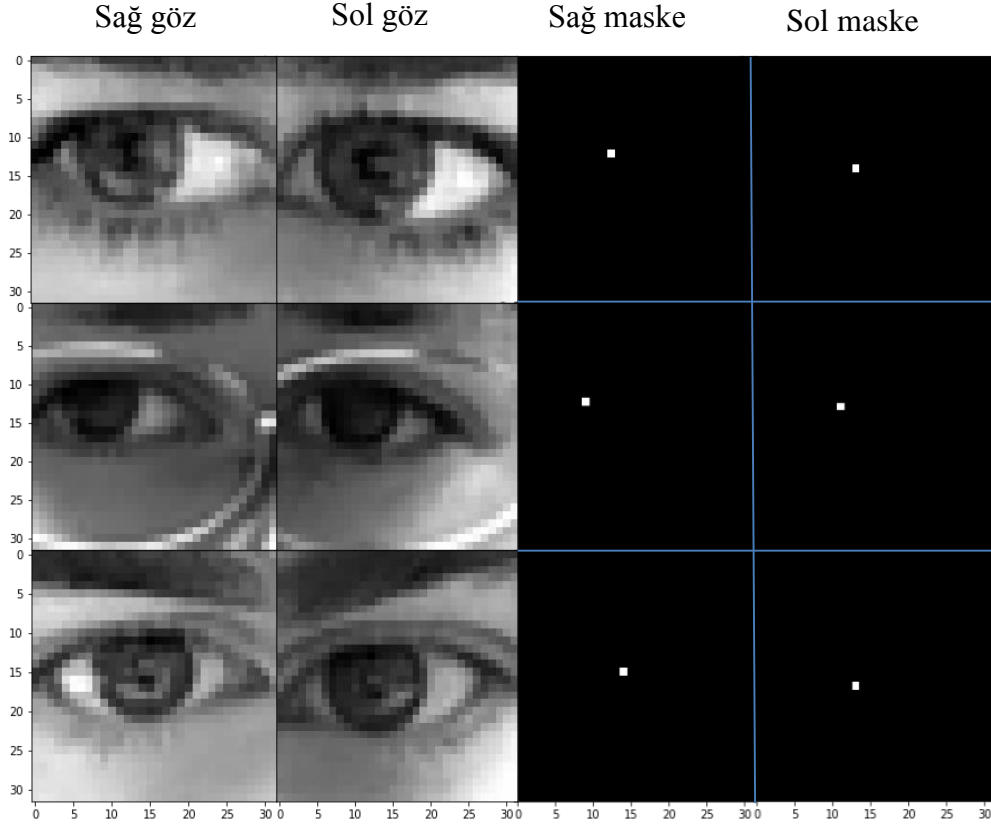
Önerilen Mini U-Net mimarisi orijinal U-Net mimarisine ince ayar yapılarak göz bebeği lokalizasyonuna uygun hale getirilmiştir. Bu mimaride giriş katmanı 32x32 uzamsal boyutlara sahip göz görüntülerine uygun hale getirilmiştir. Mimarinin kodlayıcı tarafında 4 yerine 3 adet evrişim bloğu ve her blokta evrişim işlemlerini takip eden ReLU aktivasyon fonksiyonu bulunmaktadır. Her blok sonunda yine kanal sayısı iki katına çıkarılmış ve maksimum havuzlama işlemi uygulanarak uzamsal boyut yarıya düşürülmüştür. Evrişim işlemlerinde uzamsal boyutlar dolgu (padding) kullanılarak sabit tutulmuştur. U-Net mimarisindeki darboğaz (bottleneck) katmanı 1024 kanal ile biterken önerilen mimaride 256 kanal ile sonlanmıştır. Kod çözücü tarafta ise yine 4 yerine 3 adet evrişim bloğu kullanılmış. Her blok girişinde yine üst örnekleme yapılarak uzamsal boyut iki katına çıkarılmıştır. Atlama bağlantısı ile gelen tensör üst örnekleme yapılarak gelen tensör ile birleştirilmiştir. Kanal sayısı her blok sonunda evrişim işlemleri ile yarıya düşürülmüştür.

Kod çözücünün son evrişim bloğunun son katmanına (1x1) evrişim uygulanmıştır. Orijinal mimariden farklı olarak son çıkış aktivasyon fonksiyonu olarak lokalizasyonda iyi sonuçlar vermesi nedeniyle “swish” olarak değiştirilmiştir. Segmentasyon/Bölütleme için kanal sayısı 1’e düşürülerek 32x32x1 uzamsal boyutlara sahip U-Net mimarisinin aksine 2 yerine 1 sınıflı çıkış görüntüsü elde edilmiştir. Şekil 3.4’te önerilen Mini U-Net mimarisi gösterilmiştir.



Şekil 3.4: Mini U-Net ağ yapısı

Eğitim örnekleri GI4E veri seti içinde yer alan 1236 adet yüz görüntüsünden elde edilmiştir. Bu görüntüler sağ ve sol olmak üzere 3 kanallı 2472 adet göz görüntüsü GI4E veri setinde köşe koordinat etiketleri kullanılarak kırılmıştır. Kırpma işlemi OpenCV kütüphanesi aracılığı ile yapılmış olup gri ölçekte 32x32 piksel boyutunda görüntüler elde edilmiştir. Elde edilen göz görüntüleri %80’ eğitim ve %20 test aşamasında kullanılmak üzere ayrılmıştır. Mini U-Net ile göz bebeği merkezi lokalizasyonunun gerçekleştirilebilmesi için ağın eğitimi için gerekli maskelenmiş göz merkezi görüntüleri ile beslenmesi gerekmektedir. Bunun için göz 32x32 uzamsal boyutlarındaki eğitim ve test görüntülerine karşılık gelen yer işaretleri Şekil 3.5’te gösterildiği gibi gözün merkez noktası beyaz ve kalan kısım ise siyah olarak maskelenmiştir.



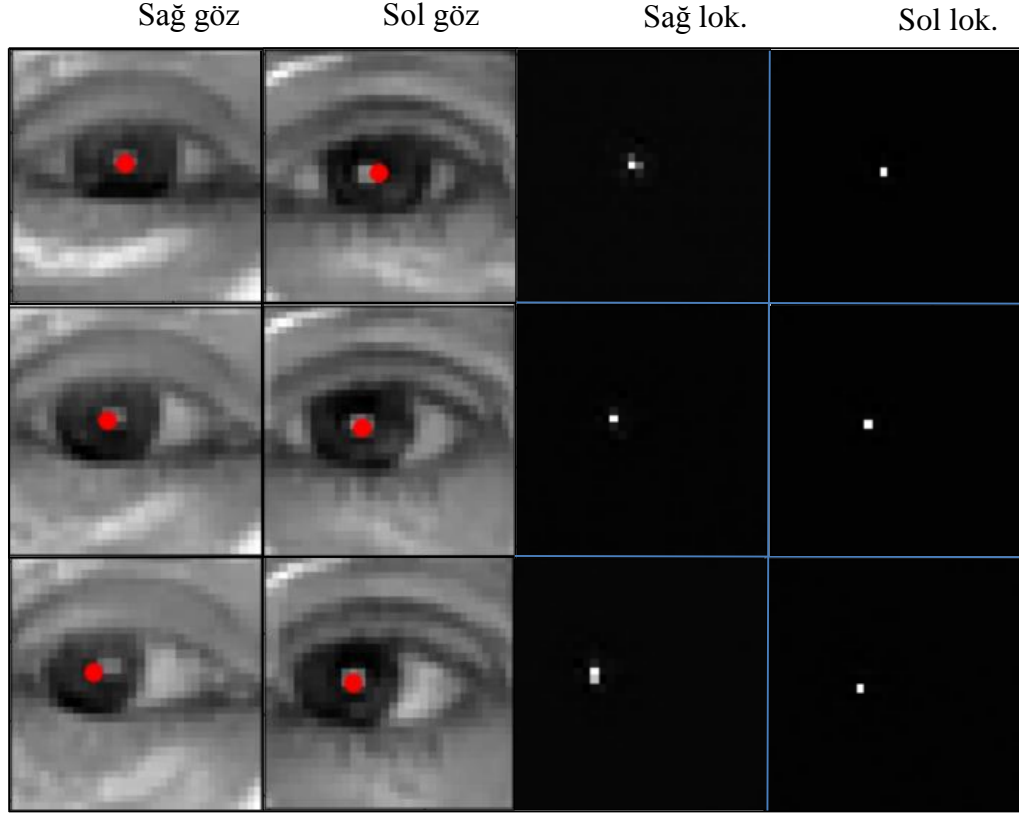
Şekil 3.5: Göz görüntülerinin maskelenmesi

Eğitim ve etiket bilgileri hazırlanan ağıın eğitim aşamasında öğrenmenin gerçekleşmesi için kayıp fonksiyonu ve optimizasyon algoritmalarına ihtiyaç vardır. Kayıp fonksiyonu olarak tahmin edilen ve gerçek değerler arasındaki ortalama farkı hesaplayan Ortalama Kare Hata (OKH-MSE) kullanılmıştır. Optimizasyon algoritması olarak da probleme uygun en iyi optimizasyonu gerçekleştiren “Adam” optimizasyon algoritması kullanılmıştır. OKH de amaç iki resmin benzerliğini ortaya koymaktır. Dolayısıyla iki resim arasındaki piksel değerleri arasındaki farkın karesinin tüm piksel sayısına bölünmesiyle hata değeri elde edilmiştir. OKH eşitliği Denklem 3.1’de verilmiştir.

$$OKH = \frac{1}{W*H} \sum_{i=1}^W \sum_{j=1}^H (y_{i,j} - \hat{y}_{i,j})^2 \quad (3.1)$$

Denklemden 1’de verilen W ve H görüntünün boyutları, $y_{i,j}$ ilgili pikselin gerçek değeri, $\hat{y}_{i,j}$ ise tahmin edilen ilgili piksel değerini ifade etmektedir. Mini U-Net ağı 1972 adet eğitim görüntüsü ve 1972 adet maskelenmiş göz görüntüsü ile eğitilmiştir. Eğitim 250 çevrimde minimum OKH değerine ulaşmıştır. Eğitim sonunda daha önce kullanılmamış 500 adet test verisi ile ağıın performansı test edilmiştir. Test görüntülerinin Mini U-Net ağına

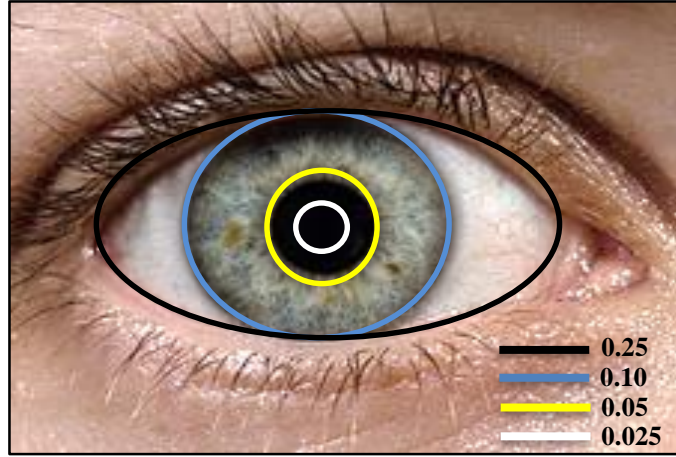
verilmesiyle göz bebeği merkezi lokalizasyonu gerçekleştirilmiştir. Ağ tarafından gerçekleştirilen segmentasyon sonunda tahmin edilen çıkış görüntüsü üzerinde OpenCv kütüphanesi aracılığıyla en yoğun değere sahip piksel, göz merkezi olarak kabul edilmiştir. Şekil 3.6’da önerilen ağ mimarisinin bazı test verileri üzerindeki lokalizasyon tahminlerinin, gerçek görüntüler üzerindeki konumları kırmızı nokta ile işaretlenmiştir.



Şekil 3.6: Test verisi lokalizasyon tahminleri

3.4 Deneysel Sonuçlar

Eğitilen Mini U-Net ağının son teknoloji algoritmalar ile performans karşılaştırılmasının yapılması için göz bebeği lokalizasyonunda yaygın olarak kullanılan bir değerlendirme kriteri olan Jesorsky ve diğ. [108] tarafından normalleştirilmiş hata değeri kullanılmıştır. Normalleştirilmiş hata değeri 0.025’e eşit veya küçükse tahmin edilen göz merkezi, gerçek göz merkezine en yakın konum, 0.05’e eşit veya küçükse tahmin edilen göz merkezi konumu, gerçek göz bebeğinin sınırları içerisinde, 0.10’a eşit veya küçükse iris sınırları içerisinde ve son olarak 0.25’e eşit veya küçükse göz merkezi ile göz köşesi arasındaki mesafede bir konuma işaret etmektedir. Bu durum Şekil 3.7’de gösterilmiştir.

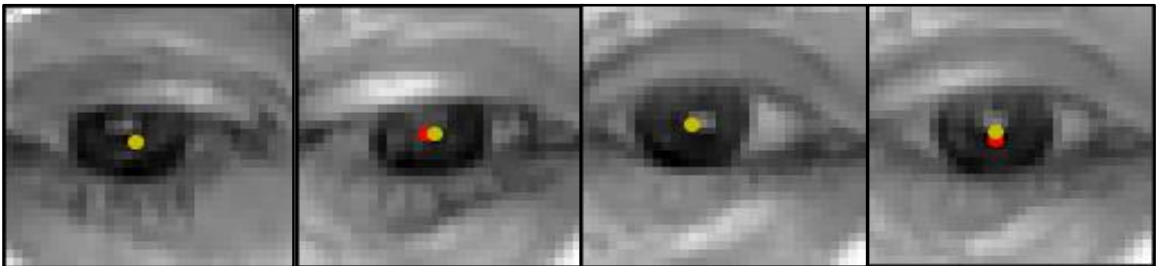


Şekil 3.7: Normalleştirilmiş hata değeri sınırları

Normalleştirilmiş hata eşitliği Denklem 3.2 ve 3.3'te verilmiştir. Normalleştirilmiş hata kriteri olarak göz bebeği merkezine en yakın konumu ifade eden 0.025 hatasından küçük lokalizasyon tahminlerinde Mini U-Net ağından %98.40'lık bir doğruluk elde edilmiştir. Şekil 3.8'de verilen test görüntülerindeki kırmızı noktalar gerçek göz bebeği merkezi konumunu sarı noktalar ise önerilen ağı tahmin ettiği göz bebeği lokalizasyonunu ifade etmektedir.

$$hata_{mak} = \frac{\max(m_{sol}, m_{sağ})}{SSGM} \quad (3.2)$$

$$SSGM = \sqrt{(x_{i,sol} - x_{i,sağ})^2 + (y_{i,sol} - y_{i,sağ})^2} \quad (3.3)$$



Şekil 3.8: Gerçek ve tahmin lokalizasyonlarının karşılaştırılması

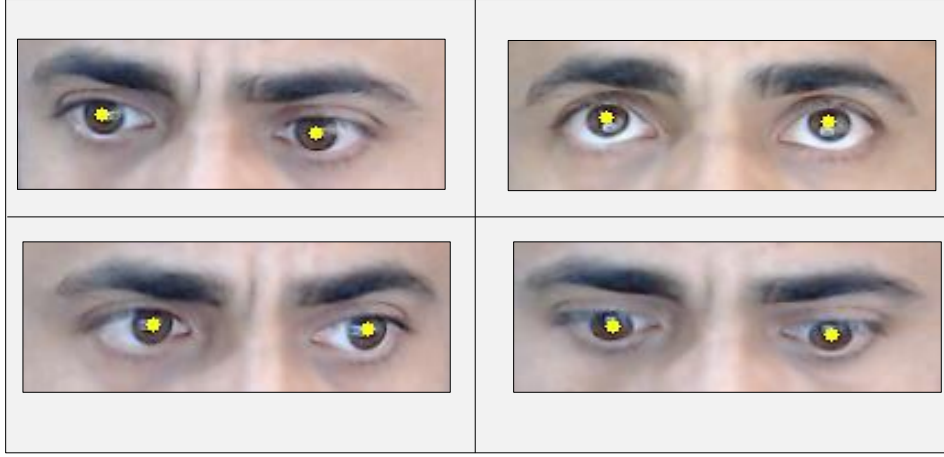
Denklem 3.2'de yer alan $hata_{mak}$ maksimum hata ölçüsünü, m_{sol} tahmin edilen sol göz bebeği merkezi konumu ile gerçek sol göz merkezi konumu arasındaki mesafeyi, $m_{sağ}$ tahmin edilen sağ göz bebeği merkezi konumu ile gerçek sağ göz merkezi konumu

arasındaki mesafeyi, *SSGM* sol göz bebeği merkezi ile sağ göz bebeği merkezi arasındaki gerçek mesafeyi ifade etmektedir. Denklem 3.3'te yer alan *SSGM* denkleminde, $x_{i,sol}$, $x_{i,sağ}$ sırasıyla sol ve sağ göz bebeği merkezlerinin x koordinatlarını, $y_{i,sol}$, $y_{i,sağ}$ sırasıyla sol ve sağ göz bebeği merkezlerinin y koordinatlarını ifade etmektedir. GI4E veri seti ile yapılan göz bebeği merkezi lokalizasyonunun son teknoloji algoritmalarla yapılan doğruluk karşılaştırması Çizelge 3.1'de verilmiştir.

Çizelge 3.1: GI4E veri seti performans karşılaştırması

Yöntem	$hata_{max} \leq 0.025$	$hata_{max} \leq 0.05$	$hata_{max} \leq 0.1$	$hata_{max} \leq 0.25$
Zhang ve diğ. [109]	-----	% 97.90	% 99.60	% 99.92
Gou ve diğ. [110]	-----	% 98.20	% 99.80	% 99.80
Gou ve diğ. [66]	-----	% 94.20	% 99.10	% 99.80
Levinshtein ve diğ. [111]	% 88.34	% 99.27	% 99.92	% 100
Cai ve diğ. [102]	% 85.70	% 99.50	-----	-----
Xiao ve diğ. [112]	-----	% 97.90	% 100	% 100
Kitazumi ve Nakazawa [103]	% 96.28	% 98.62	% 98.95	-----
Choi ve diğ. [104]	% 90.40	% 99.60	-----	-----
Xia ve diğ. [113]	-----	% 99.10	% 100	% 100
Kim ve diğ. [114]	% 79.50	% 99.30	% 99.90	-----
Lee ve diğ. [115]	-----	% 99.84	% 99.84	% 100
Larumbe-Bergera ve diğ. [105]	% 98.46	% 100	% 100	-----
Önerilen mimari (Mini U-Net)	% 98.40	% 99.40	% 99.60	%99.80

Çizelge 3.1 incelendiğinde 0.1 hata ve 0.25 normalleştirilmiş hata karşılaştırmasında algoritmaların başa baş sonuçlar sergilediği görülmektedir. Fakat gerçek göz bebeği merkezinin konumuna en yakın hata oranı olan 0.025 ve 0.05 durumlarında önerilen Mini U-Net ağının iyi bir başarımla sergilediği görülmektedir. Şekil 3.9'da önerilen ağın kameradan alınan anlık görüntülerden göz bebeği merkezi lokalizasyonu görüntüleri verilmiştir.



Şekil 3.9: Gerçek zamanlı gözbebeği lokalizasyonu

Bu çalışmada göz bebeği merkezinin lokalizasyonunda, görüntü bölütlemeye başarısını kanıtlamış olan derin öğrenme tabanlı U-Net mimarisi baz alınarak daha düşük seviyeli bir Mini U-Net algoritması önerilmiştir. Önerilen algoritmaya girdi olarak verilen GI4E veri setinin göz görüntülerinden göz bebeği merkezi koordinatları 0.025 normleştirilmiş hata kriterine göre %98.40'lık bir doğrulukla tahmin edilmektedir. Sonuçlar literatürde önerilen diğer algoritmalarla karşılaştırılmıştır. Karşılaştırma sonunda önerilen algoritmanın başarısı açık bir şekilde görülmektedir. Önerilen algoritma gerçek zamanlı bakış takip uygulamalarında göz merkezinin doğru bir şekilde tespitinde kullanılabilir.

4. GÖRÜNTÜ İŞLEMEYE DAYALI GERÇEK ZAMANLI BAKIŞ TAKİBİ

Bu bölümde bilgisayar ekranına yerleşik bir web kamerası ile düşük maliyetli bir bakış takip sistemi geliştirilmiştir. Bu sistem, görüntü işleme tekniklerini kullanan gerçek zamanlı bir sistemdir. Sistemin kamerasından alınan anlık yüz görüntülerine iris merkez koordinatlarının tespiti için makine öğrenimi ve bir dizi morfolojik görüntü işleme teknikleri uygulanmıştır. Daha sonra kamera ve ekran koordinat ilişkisinin haritalanması için 9 noktalı bir kalibrasyon prosedürü uygulanmıştır. Kalibrasyon sonrası irisin merkez koordinatlarına karşılık gelen ekran koordinatları ara değerlendirme yöntemi ile hesaplanmış ve ekrana bakan kullanıcının ekranda baktığı koordinatlar tespit edilmiştir.

Bu bölümde ilk olarak literatürde görüntü işlem algoritmaları ile geliştirilmiş bakış takip çalışmalarına yer verilmiştir. Daha sonra geliştirilen bakış takip sisteminin çalışma prosedürü ve kullanılan algoritmalar anlatılmıştır. Son olarak geliştirilen bakış takip sisteminin doğruluğu, uygulanabilirliği test edilmiş ve deneysel sonuçlara yer verilmiştir.

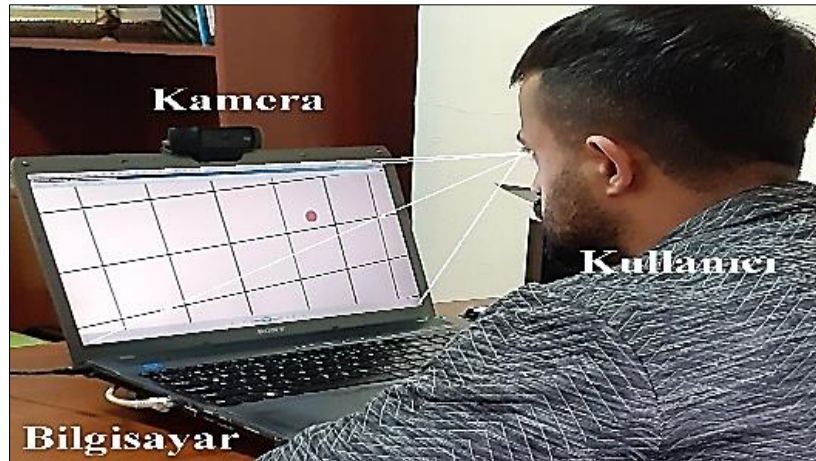
4.1 İlgili Çalışmalar

Fu ve Yang [116] dizüstü bilgisayara yerleşik kamera ile yaptıkları çalışmada göz tespiti yapıldıktan sonra iris merkezini tespit etmek için üç adımlı bir yöntem uygulamışlardır. Bu yöntem sırasıyla görüntü iyileştirme, kenar algılama ve Hough daire algılama adımlarından oluşmaktadır. Elde edilen iris merkezi referans alınarak gözün hareket edebileceği sağ ve en sol sınırlar belirlenmiştir. Buna göre elde edilen sınırlar ekran sınırları ile kalibre edilerek sınırlı bir bakış takip sistemi önerilmiştir. Bakış takibini test etmek için ekran üzerinde bulunan bir resim sağa ve sola göz hareketleri yardımıyla kaydırılmıştır. Utaminingrum ve diğ. [117] yaptıkları çalışmada engelli bireylerin gözlerini kullanarak erişilebilirliklerini sağlamak için bir yöntem geliştirmişlerdir. Geliştirdikleri yöntemde bireyin hangi yöne baktığı tespit edilmektedir. Yöntemde ilk olarak Haar Cascade algoritması kullanılarak göz bölgeleri tespit edilmiştir. Elde edilen göz bölgelerine bir dizi ön işlem uygulanmıştır. Göz bölgeleri sırasıyla gri ton görüntüye çevirme, yumuşatma, uyarlanabilir eşikleme, morfolojik açma-kapama gibi işlemler sonunda gözdeki irisin tespiti

için uygun hale getirilmiştir. Ardından, iris bölgesi tespit edilen göz bölgesinin merkezi konumu dikkate alınarak öne, sağa ve sola bakış tespiti yapılmıştır. Tao ve Pang [118] dört adımdan oluşan bir göz izleme sistemi geliştirmişlerdir. Göz izleme sistemi yüz tespiti, göz tespiti, göz izleme ve fare kontrolünden oluşmaktadır. Yüz tespiti için yüzün ten rengi baz alınarak 2D Gauss modeli kullanılarak yüz bölgesi arka plandan ayrılmıştır. Elde edilen yüz bölgesinden göz bölgelerinin tespiti için aşamalı eşikleme kullanılmıştır. Daha sonra elde edilen bölütlenmiş görüntüdeki gürültüleri kaldırmak için sırasıyla genişleme ve aşındırma işlemleri kullanılarak iris bölgesi ve iris merkezi tespit edilmiştir. Göz hareketlerinin takibi için ortalama kaydırma algoritması kullanılarak bakış pozisyonu tahmin edilmiştir. Park ve Kim [119] yaptıkları çalışmada yüz bölgelerinden göz bölgelerinin elde edilmesinde Dlib yüz dedektörünü kullanmışlardır. Elde ettikleri göz görüntülerine kontrast, yumuşatma ve orta boyutlu filtreler gibi morfolojik işlemler uygulayarak iris bölgesinin tespitini kolaylaştırmışlardır. Morfolojik işlemlerden sonra elde edilen görüntüye Hough Dönüşümü uygulanarak iris dairesi tespit edilmiştir. Elde edilen irisin merkez noktası baz alınarak yapılan kalibrasyon işleminden sonra göz bakış takibi gerçekleştirilmiştir.

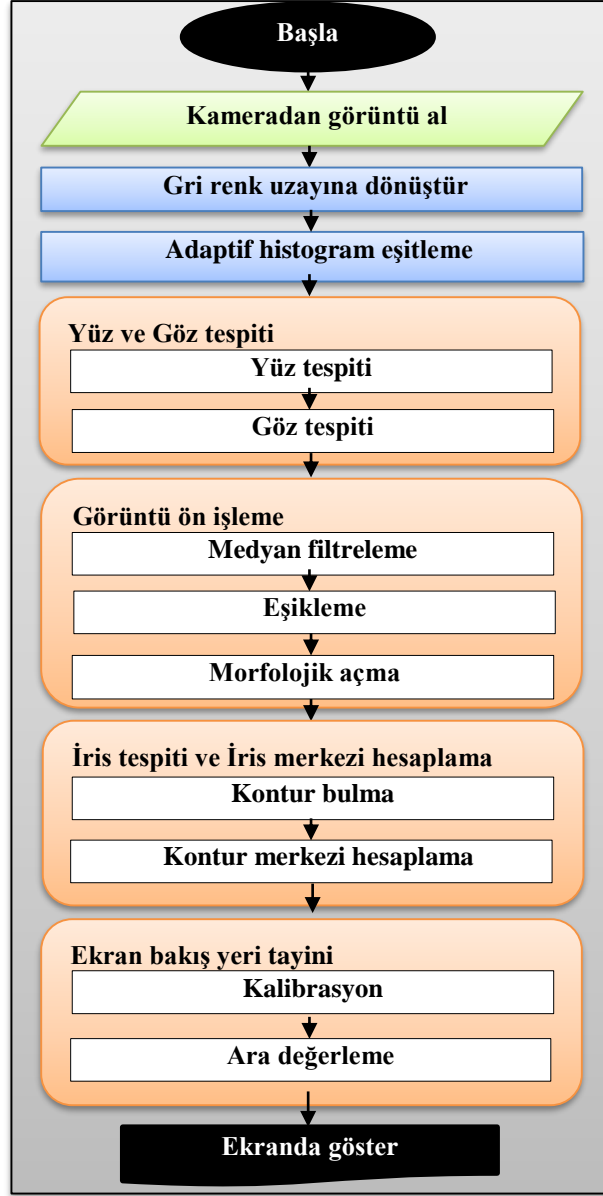
4.2 Bakış Takip Sistemi ve Algoritma

Geliştirilen bakış takip sistemi, dizüstü bilgisayara yerleşik bir web kamerası (Logitech C920) ve Python programlama yazılımı ile geliştirilmiş bir bakış takip arayüzünden oluşmaktadır. Sistem kamerasından alınan gerçek zamanlı görüntü kareleri morfolojik görüntü işleme teknikleri ile işlenerek kullanıcının ekran üzerindeki bakış noktasını tespit edilmiştir. Geliştirilen bakış takip sisteminin çalışma düzeni Şekil 4.1’de gösterilmiştir.



Şekil 4.1: Bakış takip sistem kurulumu

Bakış takip sistemi ile kullanıcın ekran üzerindeki bakış koordinatlarının belirlenmesi 4 aşamada gerçekleşmektedir. Bu aşamalar sırasıyla Yüz ve Göz Tespiti, İris ve İris Merkezinin Tespiti, Kalibrasyon, Ekran Bakış Yeri Tespiti olarak özetlenebilir. Bakış takip sistemi algoritmasının akış şeması Şekil 4.2’de verilmiştir.

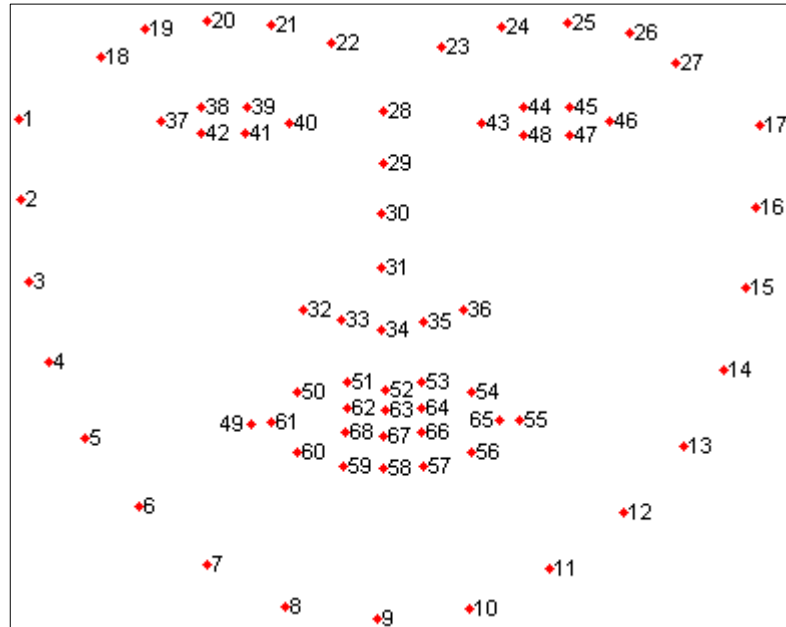


Şekil 4.2: Bakış takibi sistemi akış şeması

4.2.1 Yüz ve göz tespiti

Sistem kamerasından elde edilen anlık görüntülere uygulanacak morfolojik işlemlerin kullanım kolaylığı için orijinal görüntüler gri renk tonuna çevrilir. Daha sonra yüz ve göz tespiti aşamasında tespit edilen görüntülerde daha isabetli sonuçlar alabilmek için yüz

görüntülerine ön işlem prosedürleri uygulanır. İlk uygulanan prosedür kontrast sınırlı adaptif histogram eşitlemedir. Standart histogram, görüntünün tamamı üzerinde bir eşitleme gerçekleştirirken adaptif histogram eşitleme, ideal göz görüntüsünün elde edilebilmesi için yüzü belli ölçülerde bölgelere ayırarak bu işlemi gerçekleştirir. Ayrılan her bölge için belirli aralıklarda toplanmış piksel yoğunluk değerlerinin 0-256 arasında dağıtılmasıyla kontrast artırma olayı gerçekleştirilmiş olur [120]. Bu eşitleme yöntemi görüntüdeki nesnelere belirginleşmesini ve diğer nesnelere ayrılmasını sağlayarak nesne tanıma ve morfolojik görüntü işleme tekniklerinin verimli bir şekilde uygulanmasına zemin hazırlar. Bu aşamadan sonra elde edilen görüntülerden yüz ve gözlerin tespiti için bir nesne tanıma yöntemine ihtiyaç duyulmuştur. Bunun için yüzdeki ağız, burun, göz gibi uzuvların dönüm noktalarının tespitinde başarılı bir makine öğrenme çerçevesi olan Dlib kullanılmıştır. Dlib kütüphanesinin eğitilmiş yüz işaret detektörü sayesinde yüzdeki kaşlar, gözler, ağız ve burun bölgelerinin sınır koordinatları 68 koordinat ile tespit edilmiştir. Şekil 4.3'te yüz dönüm noktalarına karşılık gelen 68 yüz işareti gösterilmiştir. Şekil 4.4'te yüz işaret detektörünün bir görüntüye uygulanması gösterilmiştir. Bu aşamadan sonra Şekil 4.3'teki sağ göz sınırını belirleyen 43-48 arası dönüm noktası koordinatları kullanılarak açık kaynak bilgisayarlı görme kütüphanesi olan OpenCV'nin maskeleyme aracıyla göz sınırları, görüntünün geriye kalanından ayrılmıştır.



Şekil 4.3: 68 adet dönüm noktası gösterimi [93]



Şekil 4.4: Dönüm noktalarının yüze uygulanması

4.2.2 İris ve iris merkezinin tespiti

Göz tespiti yapıldıktan sonra iris bölgesinin belirlenebilmesi için görüntüden çıkarılan göz bölgesine bazı morfolojik görüntü işleme tekniklerinin uygulanması gerekmektedir. Böylelikle irisin, gözün geriye kalan kısmından ayrılması mümkün olacaktır. Genellikle gri ton çevrimi yapılmış görüntülere uygulanan morfolojik görüntü işleme tekniklerindeki temel amaç görüntüdeki nesnelere ve sınırlarını belirginleştirerek daha iyi ayırt edilmesini sağlamaktır. İris bölgesini göz bölgesinden ayırmak ve iris bölgesindeki ışık yansımalarını ortadan kaldırmak için sırasıyla medyan filtreleme, eşikleme ve morfolojik açma (aşındırma, yayma) görüntü ön işlem teknikleri uygulanmıştır.

4.2.2.1 Medyan filtreleme

Görüntünün bulanıklaştırılmasında ve görüntü üzerindeki gürültülerin giderilmesinde kullanılan doğrusal olmayan filtreleme işlemidir [121]. Bu işlem ile görüntünün her piksel değeri, pikselin etrafındaki komşu piksellerin değerleri sıralandıktan sonra elde edilen sıralamanın ortanca değerinin mevcut piksele aktarılmasıyla elde edilir. Böylelikle görüntüde daha yumuşak geçişler sağlanmış olur.

4.2.2.2 Eşikleme

Eşikleme, görüntüdeki her piksele aynı eşik değerinin uygulanmasıyla ikili (binary) formatta bir görüntü elde edilmesi işlemidir. Bu işlem, görüntünün her piksel değeri için


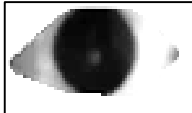

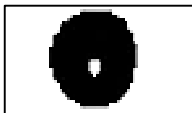


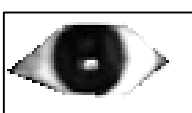
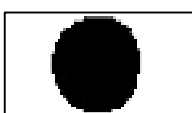
uygulanır. Eğer piksel değeri belirlenen eşik değerden küçükse 0, değilse 1 değeri verilerek görüntü oluşturulur [122]. İkili görüntü üzerinde morfolojik işlemler daha kolay yapılmaktadır.

4.2.2.3 Morfolojik açma

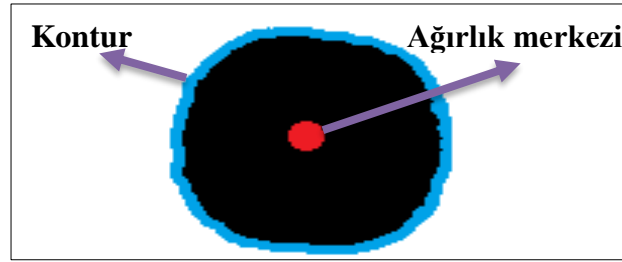
Morfolojik açma, aşındırma (erosion) ve genişletme (dilation) işlemlerinin görüntüye sırasıyla uygulanmasıyla gerçekleştirilen bir görüntü işleme yöntemidir. Her iki yöntemde de görüntü üzerinde $n \times n$ 'lik bir çekirdek matris kaydırılarak bir konvolüsyon işlemi gerçekleştirilir. Aşındırma işlemi, ilgili nesnenin (beyaz) küçülmesi veya sınırlarının aşındırılması ile sonuçlanır. Böylelikle iristeki yansıma (beyaz) ortadan kaldırılır. Genişletme işleminde ise aşındırmanın aksine bir işlem gerçekleştirilerek ilgili nesnenin (beyaz) sınırlarının genişletilmesi ile sonuçlanır. Morfolojik açma formülü Denklem 4.1'de verilmiştir. Denklemdeki A işlenecek görüntü, B ise aşındırma ve genişletme işlemlerini gerçekleştirecek yapı elemanıdır [122]. Ayrıca Çizelge 4.1'de görüntüye uygulanan tüm görüntü işleme adımları sırasıyla uygulamalı olarak gösterilmiştir.

$$A \circ B = (A \ominus B) \oplus B \quad (4.1)$$

Çizelge 4.1: Görüntü işleme adımları

Görüntü işleme adımları					
No	Görüntü	İşlem	No	Görüntü	İşlem
1		Ham görüntü	5		Medyan filtreleme
2		Gri ton	6		Eşikleme
3		Histogram	7		Aşındırma
4		Maskeleme	8		Genişletme

Morfolojik açma işleminden sonra elde edilen işlenmiş son görüntüden iris ve iris merkezinin tespiti gerçekleştirilebilir. İrisin merkezi, göz bebeği merkezinin tahmini yerini göstermektedir. Dolayısıyla iris merkezinin tespit edilebilmesi için öncelikle iris hatlarının belirlenmesi gerekmektedir. Bu aşamada iris tespiti için kontur bulma tekniği kullanılmıştır. Bu teknik, aynı renk veya yoğunluğa sahip tüm kesintisiz noktaların birleştirildiği eğri olarak ifade edilmektedir [123]. Konturlar nesne algılama ve şekil analizi için uygun bir araçtır. Kontur bulma işleminde OpenCV kontur bulma aracı kullanılmıştır. Şekil 4.5'te kontur uygulaması ve ağırlık merkezi gösterilmiştir.



Şekil 4.5: Kontur uygulaması ve ağırlık merkezi

Kontur bulma işlemiyle tespit edilen iris bölgesi ışık yansımalarından dolayı her zaman düzgün bir daire şeklinde olmayabilir. Bu yüzden iris merkezinin tespiti için düzenli ve düzensiz şekillerin merkezi (centroid) koordinatlarını veren Denklem 4.2-4.6 arasındaki ağırlık merkezi geometrik formülleri kullanılmıştır. Denklemlerde yer alan $I(x,y)$ konturun (x, y) koordinatındaki piksel değerini, " M " ifadeleri ise sırasıyla merkez, x ve y eksen moment değerlerini ifade etmektedir [124].

$$M_{00} = \sum_x \sum_y I(x,y) \quad (4.2)$$

$$M_{10} = \sum_x \sum_y x \cdot I(x,y) \quad (4.3)$$

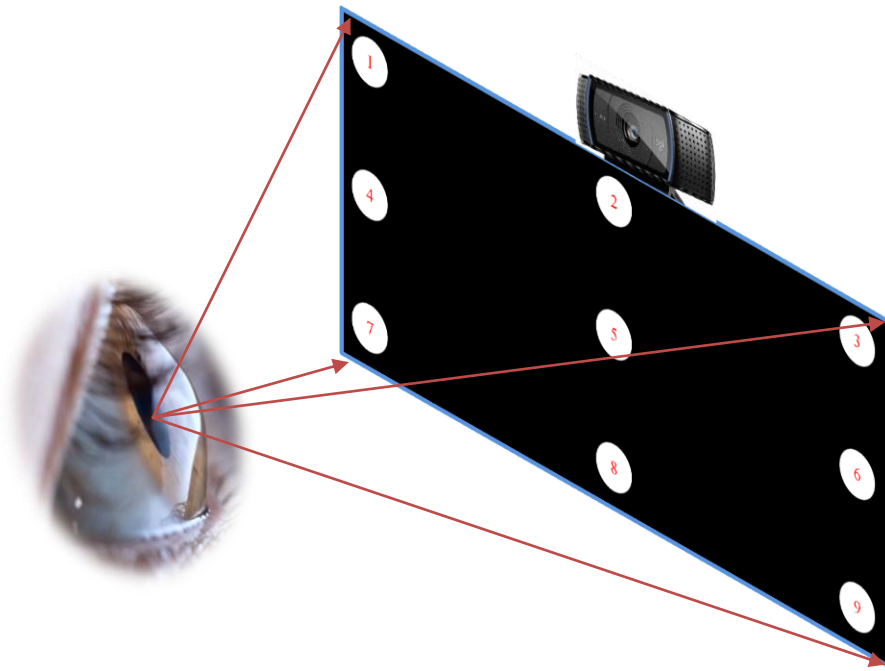
$$M_{01} = \sum_x \sum_y y \cdot I(x,y) \quad (4.4)$$

$$x_{\text{centroid}} = \frac{M_{10}}{M_{00}} \quad (4.5)$$

$$y_{\text{centroid}} = \frac{M_{01}}{M_{00}} \quad (4.6)$$

4.2.3 Kalibrasyon prosedürü

Kameradan alınan göz hareketlerinin bilgisayarın ekran boyutları ile sınırlandırılması için görüntü-ekran kalibrasyon prosedürünün uygulanması gerekmektedir. Bu prosedür ile görüntü-ekran haritalama ilişkisi gerçekleştirilir. Bu prosedürde 9 nokta kalibrasyonu uygulanmıştır. Bu prosedür ile kullanıcı sabit baş pozisyonuyla bilgisayar ekranında koordinatları önceden belirlenmiş 9 kalibrasyon noktasına sırayla odaklanması istenmiştir. Kullanıcı bu noktalara baktığında kamera görüntüsündeki iris merkezi koordinatları (x, y) , kalibrasyon koordinatları (X, Y) ile ilişkilendirilmiştir. Kalibrasyon düzeni Şekil 4.6'da gösterilmiştir.



Şekil 4.6: Kalibrasyon prosedürü

4.2.4 Ekran bakış yeri tayini

Kullanıcının bilgisayar ekranındaki ilgi alanlarının doğru ve sürekli tespit edilebilmesi için kalibrasyon prosedürü sonrasında iris merkezinin kameradaki hareket alanının tüm koordinatlarının bilgisayar ekranındaki koordinat karşılıklarının bulunması gerekmektedir. Bunun gerçekleştirilebilmesi için iris merkezinin kameradaki koordinat değerlerinin ekran koordinatlarına dönüştürme işlemine ihtiyaç vardır. Bunun için mevcut veri noktalarına eğriler uydurulması ile bir noktadaki olası değeri tahmin etmeye yarayan ara değerlendirme (interpolasyon) yöntemidir [125]. Denklem 4.6 ve 4.7'de kameradaki iris

merkezi koordinatları (x, y) 'nin ekrandaki kalibrasyon koordinatları (X, Y) ile ilişkisi 2. dereceden denklem formunda verilmiştir.

$$X_{\text{ekran}} = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 \quad (4.6)$$

$$Y_{\text{ekran}} = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 \quad (4.7)$$

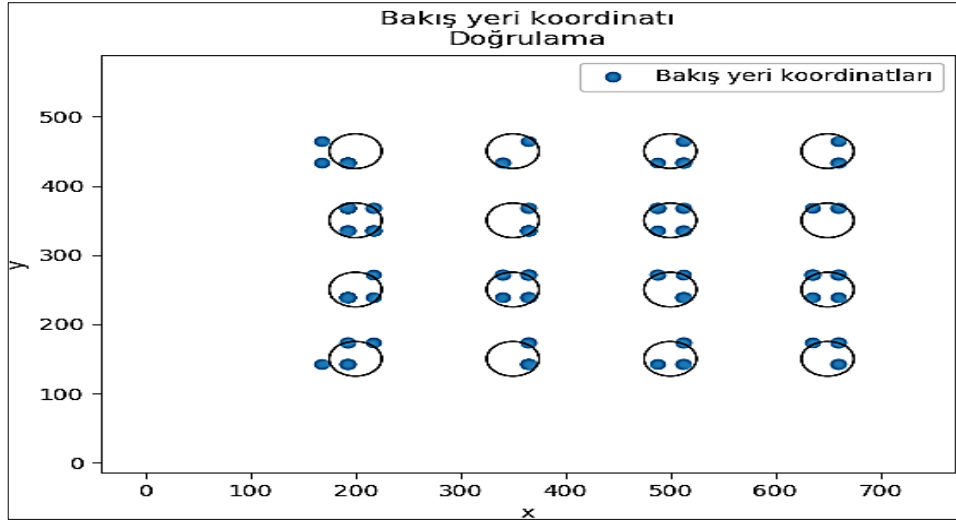
Denklem 4.6 ve 4.7'deki polinom fonksiyonlarının katsayılarının (a_n, b_n) tüm koordinatlara uygun değerlerin bulunabilmesi iris merkezi ve kalibrasyon koordinatları bir matris formuna çevrilip Denklem 4.8, 4.9'daki eşitliklere ara değerlendirme (interpolasyon) uygulanması için kullanılmıştır. Ara değerlendirme işlemi sonrası bulunan denklem sistemini sağlayan yaklaşık (a_n, b_n) değerleri 4.6 ve 4.7'deki denklemlerde yerlerine koyularak iris merkezi-ekran koordinatı dönüşümü sağlanmıştır [125–127].

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 & x_1^2 & y_1^2 \\ \vdots & & & & & \\ 1 & x_9 & y_9 & x_9y_9 & x_9^2 & y_9^2 \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_5 \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_9 \end{bmatrix} \quad (4.8)$$

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 & x_1^2 & y_1^2 \\ \vdots & & & & & \\ 1 & x_9 & y_9 & x_9y_9 & x_9^2 & y_9^2 \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_5 \end{bmatrix} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_9 \end{bmatrix} \quad (4.9)$$

4.3 Deneysel Sonuçlar

Geliştirilen bakış takip sistemi test edilerek bazı deneysel sonuçlara ulaşılmıştır. Test işleminde, 800x600 piksel çözünürlüğündeki ekranın 12 farklı koordinatına 30 piksel yarıçapında çemberler yerleştirilmiştir. Kullanıcıdan bu çemberlere sırasıyla bakması istenmiştir. Kullanıcının ekranda baktığı konumların koordinatları kayıt altına alınmıştır. Bu koordinatlar Şekil 4.7'de grafik üzerinde gösterilmiştir. Şekil 4.7 incelendiğinde kullanıcının tüm çemberlere odaklandığı görülmüştür. Bunun yanında çember yarıçaplarının 30 pikselin altında olması durumunda odaklanma doğruluğunun düştüğü görülmüştür. Odaklanma görevleri esnasında çene desteği ile baş pozisyonunun sabit olduğu göz önünde bulundurulmalıdır.



Şekil 4.7: Bakış yeri doğrulama

Şekil 4.7’de gösterilen uygulamanın performansı sabit baş pozisyonu ve belirli ortam koşulları için geçerlidir. Farklı ortam koşullarında gerçekleştirilen testlerde aydınlatma koşullarına bağlı olarak morfolojik işlemler sonucunda iris bölgesi tam olarak tespit edilemediğinden irisin merkez koordinatları da ağırlık merkezi hesaplamasına bağlı olarak belli ölçülerde hatalı tespit edilmiştir. Sonuç olarak bu bölümdeki çalışmada insan-bilgisayar etkileşiminde birçok farklı sektörde gittikçe yaygınlaşmaya başlayan gerçek zamanlı bakış takip sisteminin bir web kamerası ve görüntü işlemeye dayalı bakış takip arayüzüne dayanan bir sistem geliştirilmiştir. Geliştirilen sistemin verimli bir şekilde uygulanabilmesi için bazı kısıtlamalar mevcuttur. Bu kısıtlamalar, kamera ile kullanıcı arası mesafenin 30-40 cm olması, aydınlatma koşullarının optimum olması, kalibrasyon işleminin hatasız gerçekleştirilmesi ve en önemlisi baş pozisyonunun sabit olmasıdır. Önerilen sistem bu kısıtlamalardan dolayı profesyonel çalışmalar için henüz yeterli değildir. Kullanıcı deneyimi için daha esnek, yüksek doğrulukta çalışan bir sistem için bazı değişikliklerin yapılması gerektiği öngörülmüştür. Baş pozisyonu dikkate alınarak geliştirilecek algoritmalarla baş hareketlerinden ve morfolojik işlemlerde yansımalar dikkate alınarak geliştirilecek algoritmalarla da aydınlatma koşullarından kaynaklı hatalar belli ölçüde tolere edilebilecektir. Bunun yanında bakış takip sisteminde görünür ışık yerine kızılötesi ışığın kullanılmasıdır. Bu sayede hem göz bebeği merkezi daha doğru tespit edilecek hem de baş pozisyonu farklılıkları belli ölçüde tolere edilecektir. Tüm bu değişikliklerin yanında son yıllarda bakış takip sistemlerinde de kullanılmaya başlanan derin sinir ağlarına dayalı geliştirilecek algoritmalarla da tüm bu kısıtlamaların üstesinden gelinebileceği düşünülmektedir.

5. DERİN ÖĞRENMEYE DAYALI GERÇEK ZAMANLI BAKIŞ TAKİBİ

Günümüzde göz bakış takibi ile ilgili deneysel çalışmalarda yüksek doğrulukla çalışan ticari bakış takip cihazları kullanılmaktadır. Bu cihazlarda genellikle kızılötesi ışık ve yüksek çözünürlüklü kamera kullanılmaktadır. Bu cihazların temin edilmesi yüksek bütçeli olmalarından dolayı kısıtlıdır. Bu cihazlara alternatif, aktif bir çalışma alanı olan ESA tabanlı yaklaşımlar bakış yönünü tahmin etmede kayda değer ilerleme kaydetmiştir. Bu bölümde sadece bir web kamerası ve geliştirilen ESA tabanlı sistem ile kullanıcının web sayfalarındaki ilgi alanları belirlenmeye çalışılmıştır. Sistemde derin öğrenme tabanlı yöntem kullanılmasıyla gerçek dünya koşullarındaki baş pozunu, farklı aydınlatma ortamları, düşük çözünürlük gibi sınırlayıcı faktörler ESA tarafından birer özellik olarak öğrenilmesiyle aşmaya çalışılmıştır. Üstelik sistem, ticari ürünler gibi kalibrasyona ihtiyaç duymadan kullanılmaktadır. Sistemin performansının ölçülmesi için bir web sayfasının test oturumunda kullanıcıya verilen görevleri yerine getirmesi istenmiştir. Sonuçlar, tasarlanan sistemin web sayfalarında kullanıcıların ilgi alanlarını belirlemek için kullanılabileceğini göstermiştir.

Bu bölümde ilk olarak bakış takibi ile ilgili görünüme dayalı derin öğrenme tabanlı literatür çalışmalarına yer verilmiştir. Daha sonra iTracker[128] modelinden esinlenerek tasarlanan ESA tabanlı sistemin eğitilmesi için gerekli veri setinin oluşturulması adımları anlatılmıştır. Eğitilen ESA tabanlı sistem test aşamasını başarılı bir şekilde geçtikten sonra sistemin web sayfalarına uygulanması sağlanmıştır.

5.1 İlgili Çalışmalar

Literatürde bakış takibinde derin öğrenme çerçevesini kullanan bazı yöntemler şunlardır. Baluja ve Pomerleau [47] bireysel kullanıcılara göre özelleştirilebilen ve invazif olmayan (konforlu) Yapay Sinir Ağı (YSA) tabanlı bir bakış izleme sistemi geliştirmişlerdir. Göz görüntüsünü girdi olarak kullanan bu sistem baş hareketliliğine izin vermekte ve bakış takibini 1.5 derecelik bir doğruluk ile elde etmektedir. Sewell ve diğ. [129] engelli bireylerin

bilgisayar kullanımına olanak sağlayan ucuz maliyetli bir sistem önermişlerdir. Kullanıcının bakış konumunu kullanıcının göz görüntüsüne dayanarak tahmin etmek için bir YSA kullanılmıştır. Xu ve diğ. [130] görüntü işleme algoritmalarıyla birleştirilmiş bir sinir ağı modelleme stratejisi aracılığıyla gerçek zamanlı bakış takibinde 1.5 derece hatayla ekran üzerinde bakış tahmini gerçekleştirmişlerdir.

Göz bakış takibinde ESA tabanlı araştırmalardan bazıları ise şöyledir. Krafka ve diğ. [128] cep telefonları ve tabletler gibi ticari donanımlar üzerinde, ek sensör veya cihazlara ihtiyaç duymadan çalışan ESA tabanlı iTracker'ı geliştirmişlerdir. Bu ağın eğitimi için ilk büyük ölçekli veri kümesi olan GazeCapture'ı kullanılmışlardır. Bu veri kümesi 1450'den fazla kişiden alınan 2.5 milyon görüntü karesinden oluşmaktadır. Eğitilen model ile cep telefonlarında ve tabletlerde sırasıyla 1,71 cm ve 2,53 cm kalibrasyonsuz tahmin hatasına ulaşılmaktadır. Kim ve diğ. [131] iTracker sinir ağına benzer fakat 4 evrişim katmanlı Gazelle adlı bir ESA tabanlı model oluşturmuşlardır. Bu model ile bir iPhone kullanıcısının cihaz ekranında baktığı nokta tahmin edilmeye çalışılmıştır. Eğitiminde GazeCapture veri setinin kullanıldığı Gazelle modelinde gradyanlı histogram, normalizasyon ve maksimum havuzlama gibi parametrelerin kombinasyonları test edilmiş ve 4,85 cm hatayla sonuçlanmıştır. Diğer bir çalışmada Zhang ve diğ. [132] 15 katılımcıdan 3 aydan uzun bir sürede topladıkları 213.659 görüntüyü ve bu görüntülere ait açıklamaları içeren MPIIGaze veri setini sundular. Önerdikleri GazeNet modelinde bakış tahmini için çok modlu evrişimsel sinir ağlarını kullanmışlardır. Bu modelde normalleştirilmiş göz görüntülerinden ESA ile elde edilen özellik vektörüne tamamen bağlı katmanda baş poz vektörü eklenerek bakış yönünü tahmin ettiler. Zhang ve diğ. [133] bakış yönünü tahmin etmek için tam yüz görüntüsünü girdi olarak alan bir yöntem sunmuşlardır. Önerdikleri yöntemde yüz özelliklerini belirlemek için ESA kullanan ve farklı yüz bölgelerindeki temsili arttırmak veya baskılamak için uzamsal ağırlıklar kullanmışlardır. Hem 2B bakış noktası hem de 3B bakış vektörü tahmini gerçekleştirmişlerdir. Abdelrahman ve diğ. [134] yaptıkları çalışmada sınıflandırma ve regresyon kayıplarına dayalı L2CS-Net modelini önermişlerdir. Önerilen bu modelde tüm yüzden uzamsal bakış özelliklerini çıkarmak için RestNet-50 ağını kullanmışlardır. Bakış yönü tahmininde kullanılan ve bakış yönünü belirleyen pitch ve yaw açıları önceki çalışmaların aksine ayrı ayrı tam bağlantılı katmanda tahmin edilmeye çalışılmıştır. Her açının sınıflandırma kaybını hesaplamak için çapraz entropi kaybı uygulanmıştır. Ayrıca sınıflandırma kayıplarına ortalama karesel hata eklenerek her açı için nihai hata hesaplanmıştır. Jyoti ve Dhall [135] bakış yönü tahmini için önerdikleri yöntemde

hem geometrik hemde doku bilgisini yakalayan ve bunları daha sonra birleştiren bir yöntem önermişlerdir. Önerdikleri yöntemde yüz konum bilgilerinden çıkarılan geometrik özellikler derin sinir ağıyla eğitilmiştir. Ayrıca doku tabanlı özellikler için sağ, sol ve her iki göz ayrı ayrı evrişimsel sinir ağlarından geçirilerek eğitilmektedir. Toplam dört kanaldan elde ettikleri geometrik ve doku özelliklerini birleştirildikten sonra tamamen bağlı katmanlara vererek bakış yönünü tahmin etmişlerdir. Bir başka çalışmada Wong ve diğ. [136] mobil cihazlarda göz bakış tahmini için bir derin öğrenme modeli olan Artık Sinir Ağı (ResNet-18) kullanmışlardır. Modelin eğitimi için GazeCapture veri seti kullanılmıştır. Veri setindeki kırılmış göz verileri temizlenmiş, görüntüler üzerinde histogram normalleştirme, baş pozu ve yüz ızgarası özellikleri birleştirilerek model eğitilmiştir. Baş pozunun bakış tahmine olumlu etkisi görülmesine rağmen yüz ızgarasının kayda değer olumlu etkisinin olmadığı görülmüştür. Eğitilen modelin testler sonucunda AlexNet mimarisini kullanan derin öğrenme modeli iTracker'dan (4 nokta kalibrasyonlu 4,11 cm ortalama hata) daha iyi performans olan 3,05 cm ortalama hata elde etmişlerdir. Chen ve diğ. [137] 22 inç bilgisayar ekranı, web kamerası (Logitech C920R HD Pro) ve göz izleme (Tobii Eye Tracking 4C) cihazlarından oluşan bir sistem kurmuşlardır. Kurulan sistemle katılımcılara videolar gösterilmiştir. Veri seti, katılımcıların kamera ile görüş görüntüleri ve bakış takip cihazı ile bakış koordinatları toplanarak hazırlanmıştır. Evrişimsel katmanların, toplu normalleştirmenin ve küresel ortalama havuzlama katmanlarının sistem doğruluğu üzerindeki etkisi, bu veri setleriyle eğitilmiş ESA tabanlı bakış takip sisteminde araştırılmıştır.

Literatürde bakış takibinin web sayfalarındaki kullanımı ile ilgili araştırmalar da mevcuttur. Bu araştırmaların genelinde göz bakış takibinde 0-1 derece arası doğruluk ile çalışan Tobii, EyeLink, Gazepoint gibi firmalar tarafından üretilen ve farklı örneklem hızlarına sahip göz bakış takip cihazları kullanılmıştır [41,138,139]. Web sayfalarında göz bakış takibi ile ilgili araştırmalardan bazıları şunlardır. Slanzi ve diğ. [140] bir web sitesinin içerik ve yapısının kullanıcıların tercihlerine göre nasıl optimum yapıya getirilebileceği üzerine bir çalışma yapmıştır. Bu çalışmada web sitesinde gezinen bir kullanıcının göz bebeği genişlemesi, Sofey göz izleme sistemi aracılığıyla göz hareketi ölçümleri ve Emotiv EPOC ile EEG sinyallerinden elde edilen veriler kullanılarak kullanıcının tıklama niyetinin tahmin edilebileceğini ve kullanıcı odaklı bir oturumun yapılandırılabilceğini göstermişlerdir. Velásquez ve diğ. [141] Tobii T120 göz izleme teknoloji ile web veri madenciliğini birleştirerek web kullanıcısının bir web sitesindeki davranışları hakkında

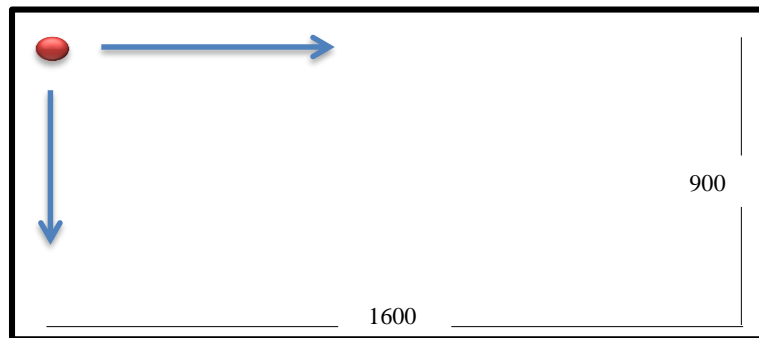
önemli bilgiler elde etmişlerdir. Böylelikle kullanıcıların gezindiği web sitelerinde en çok ilgilendikleri nesnelere tahmin etmek için anket yapmaya gerek kalmadan kullanıcının niyeti hakkında fikir sahibi olunabileceğini göstermişlerdir.

5.2 Materyal ve Metodoloji

Sağlam bir bakış takibinin gerçekleştirilebilmesi için gerçek dünya koşullarında toplanmış eğitim verilerine ihtiyaç vardır. Literatürde bakış yönü tahmini için hazırlanmış MPIIGaze [132], GazeCapture [128], Eyediap [142], UT Multiview [143], Gaze360 [144] gibi veri setleri mevcuttur. Fakat, bu tez çalışmasında önerilen ESA tabanlı bakış takibi uygulaması için python yazılımı kullanılarak yeni bir veri oluşturulmuştur. Bu veri seti üzerinde histogram eşitleme, maskeleyme, boyutlandırma ve veri temizleme gibi bir takım ön işlemler yapılmıştır. Bu ön işlemlerin amacı, veri setini ESA modelinin eğitim aşaması için uygun hale getirmektir. Tasarlanan sistemde ESA modelinin eğitimi Google Colab üzerinden gerçekleştirilmiştir [145].

5.2.1 Veri seti oluşturma

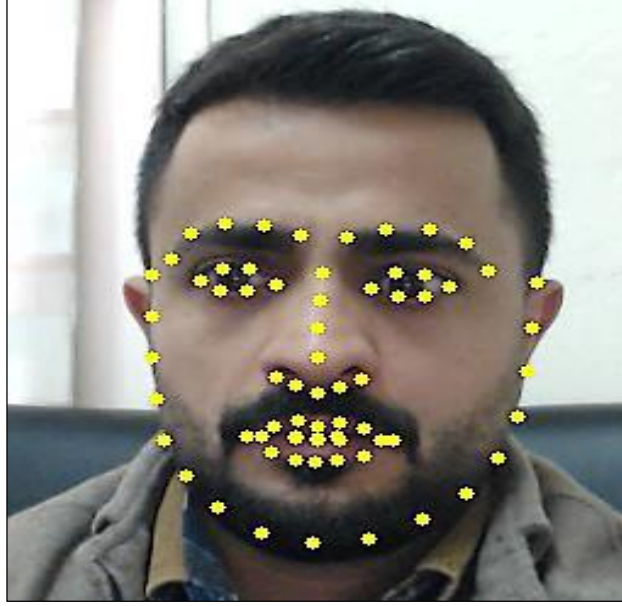
1600x900 piksellik bir ekran üzerinde python programlama dili ile hazırlanan bir uygulama ile ekrana bakan kullanıcının hareket halindeki nesneyi takip etmesi istenmiştir. Veri toplama işlemi farklı aydınlatma koşulları altında ve yatay-dikey olmak üzere 0-10 cm baş hareketi serbestliği ve kameradan 40-70 cm uzaklıklarda gerçekleştirilmiştir. Bilgisayara yerleşik kamera (Logitech C920 HD) ile eğitimde kullanılmak üzere ekrana bakan kullanıcının yüz görüntüsü, yüz konumu, sağ-sol gözler, kabaca baş pozisyonunu bilgisini veren 10 adet koordinat çifti ve nesnenin ekran üzerindeki koordinat bilgileri her nesne hareketinde anlık olarak kaydedilmiştir. Veri toplama işlemi 40 oturumda gerçekleşmiş ve toplamda her eğitim parametresi için 23040 adet kayıt yapılmıştır. Veri toplama python uygulaması Şekil 5.1’de gösterilmiştir.



Şekil 5.1: Veri toplama uygulama ekranı

5.2.2 Yüz dönüm noktaları

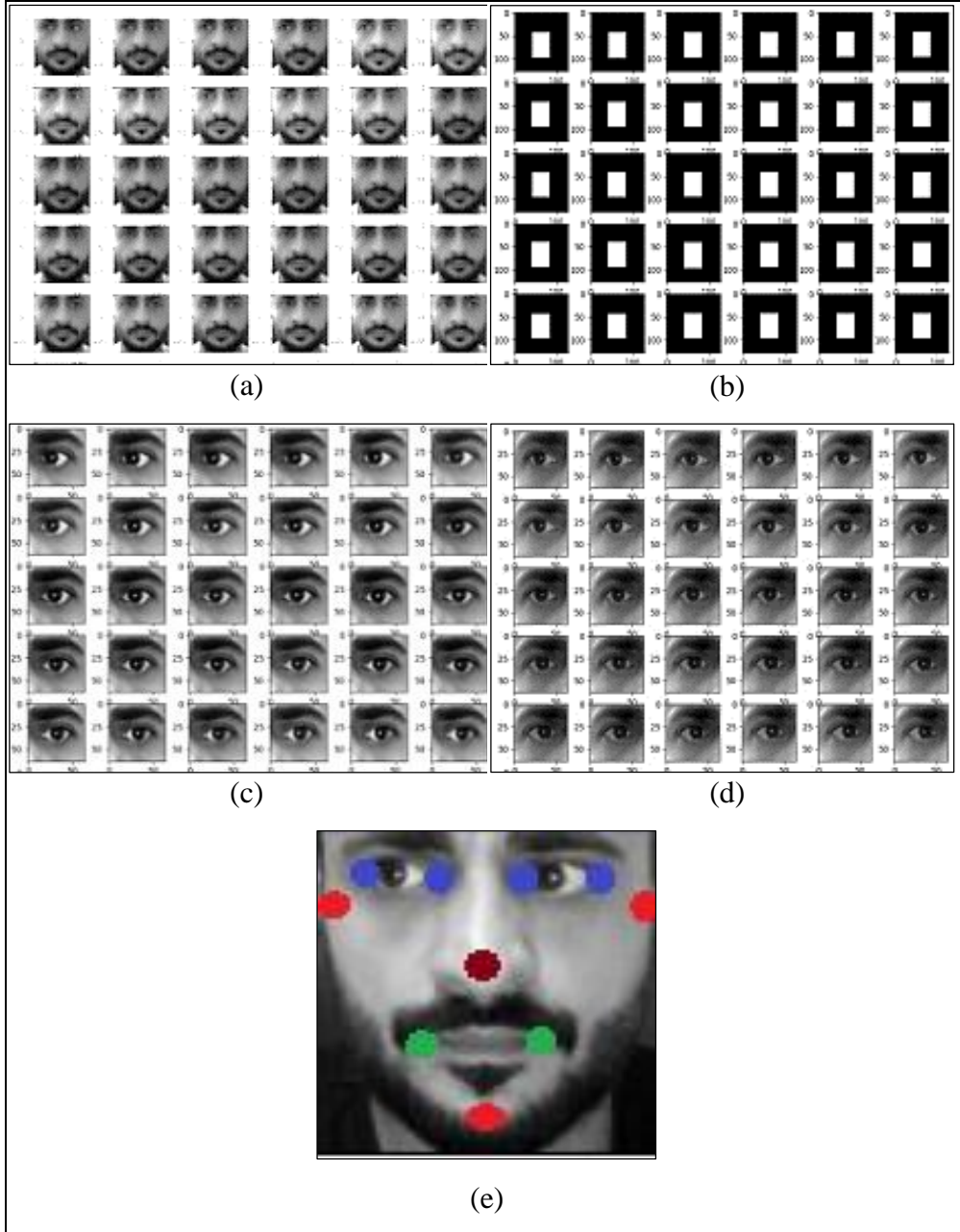
Veri seti hazırlanırken görüntülerden yüz, yüz konumu, gözler ve yüz dönüm noktalarının (10 adet) tespiti için bir nesne tanıma algoritmasına ihtiyaç duyulmuştur. Bunun için makine öğrenmesi temelli Dlib kütüphanesi kullanılmıştır. Bu kütüphanenin yüz işaret detektörü sayesinde yüz bölgesi ve yüz işaretleri tespit edilmiştir. Dlib kütüphanesinde, eğitilmiş yüz işaret detektörü yüzdeki kaş, göz, ağız, burun bölümlerinin sınır koordinatlarını 68 nokta ile tahmin etmektedir. Şekil 5.2’de yüz işaret detektörünün görüntüye uygulanması gösterilmiştir.



Şekil 5.2: Dönüm noktalarının yüze uygulanması

Bu aşamadan sonra Şekil 5.2’deki işaretçilerin koordinatları kullanılarak eğitim verileri açık kaynak bilgisayarlı görme kütüphanesi olan OpenCV maskeleye aracıyla elde edilmiştir. Bu veriler yüz, yüzün kameradaki konumu, sol göz, sağ göz ve dlib kütüphanesi yüz işaret detektörü ile 10 adet baş pozisyon koordinatları olmak üzere toplamda beş adet özelliği temsil etmektedir. Bu özelliklerden yüz bilgisi, baş ve göz hareketleriyle farklı bakış özelliklerini temsil etmektedir. Yüzün kameradaki konum bilgisi, başın kameraya olan uzaklığını ve kameraya göre konumuna karşılık gelmektedir. Sol ve sağ göz bilgileri yine odaklanılan koordinatlara ait göz yönelimlerini temsil etmektedir. Son olarak yüz bölgesinden alınan 10 adet koordinat baş pozisyonunu kabaca ifade etmektedir. Bu özelliklerin tümü beraber kullanıldığında bakış yönüne veya bakış hedefine ait spesifik özellikler olarak düşünülebilir. ESA modelinde kullanılacak OpenCV maskeleye aracı ve

dlib işaret detektörü ile elde edilen bakış koordinatlarını temsil eden özellikler (görüntü ve koordinatlar) Şekil 5.3'te gösterilmiştir.

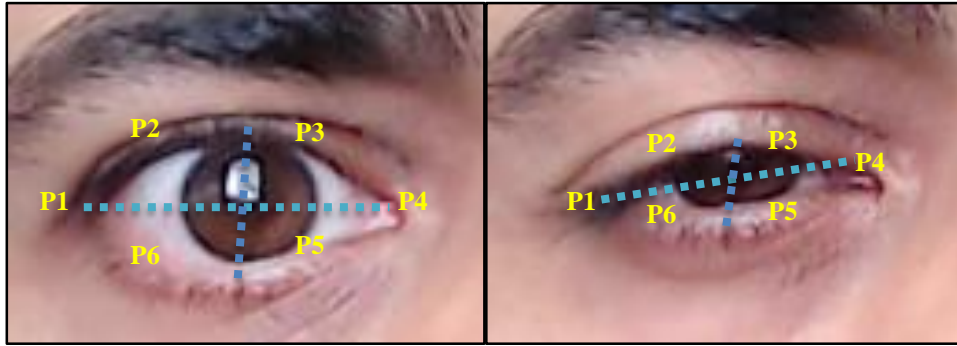


Şekil 5.3: (a) Yüz, (b) Yüz konumu, (c) Sol göz, (d) Sağ göz, (e) Baş pozisyon bilgisi

5.2.3 Veri seti temizleme ve düzenleme

ESA modelinin doğru bir şekilde eğitilebilmesi için veri sayısının yanında verinin uygunluğu da önemlidir. Veri seti oluşturma uygulamasında nesne takip edilirken gerçekleşen göz kırpmalarının ve algılanmayan yüz görüntülerinin sistemin doğru bir şekilde eğitilmesi için veri setine dahil edilmemesi gerekir. Algılanmayan görüntüler kodlar ile

engellenebilirken göz kırpılmalarının tespiti için bir algoritmaya ihtiyaç duyulmuştur. Gözdeki en-boy oranına dayanan ve göz kırpılmalarını tespit eden algoritma Denklem 5.1'deki eşitlikte verilmiştir. Denklem 5.1 ve Şekil 5.4'teki "P" işaretleri Dlib kütüphanesinin eğitilmiş yüz işaretçisi aracılığıyla tespit edilmiş göz sınırlarını ifade etmektedir. P1-P4 arasındaki mesafe gözün genişliğini, P2-P6 ve P3-P5 arasındaki mesafeler ise gözün yüksekliğini ifade eder. Bu işaretlerin koordinat değerleri kullanılarak bir hesaplama yapılmış ve bir eşik değeri belirlenmiştir [136,146]. Tasarlanan sistemin eğitimi için geçerli Göz En Boy Oranı (GEBO) eşik değeri 0,23 olarak belirlenmiştir. Bu eşik değeri üzerindeki göz görüntülerine sahip kareler veri setine dahil edilecektir [136].



Şekil 5.4: Kapalı göz tespiti

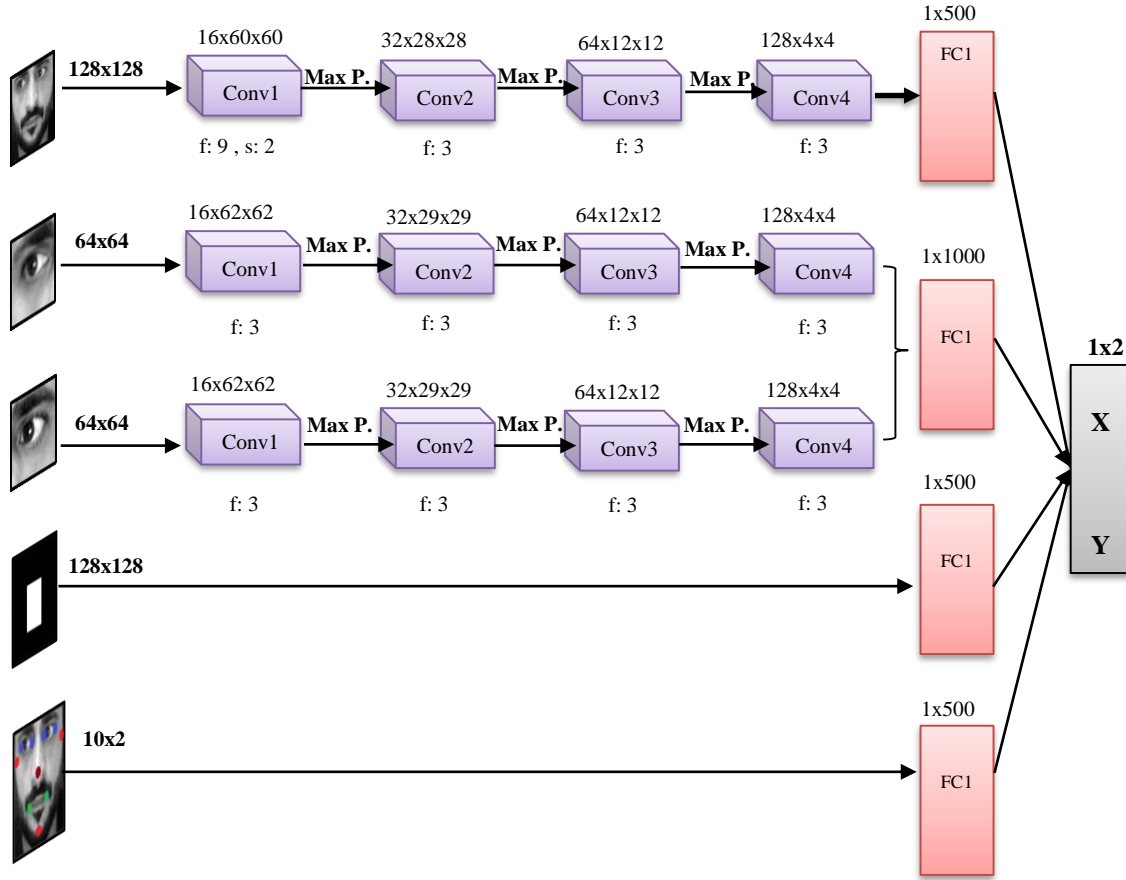
$$GEBO = \frac{|p2-p6|+|p3-p5|}{2|p1-p4|} \quad (5.1)$$

Veri toplama esnasında aydınlatma koşullarının etkisinin en aza indirilmesi için yüz üzerinde histogram eşitleme uygulanmıştır. Ayrıca elde edilen yüz ve göz bölgeleri ESA modeline uygun boyutlara dönüştürülmüştür.

5.2.4 Sistem ESA modeli ve eğitim

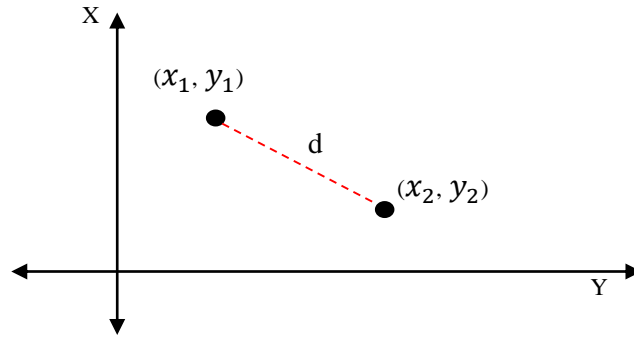
Geliştirilen sistem, iTracker[128] modelinden ilham alınarak tasarlanmış bir sistem olup öğrenim süreci sırasında göz bebeği merkezi, kappa açısı, kamera-ekran ilişkisi ve diğer geometrik parametreler ESA'nın örüntü tanımadaki yeteneğiyle dolaylı olarak öğrenilmektedir. Bakış takibi için tasarlanan sisteme, yüz görüntüsü (128x128), sol (64x64) ve sağ (64x64) gözler, yüzün kameradaki konumunu veren yüz maskesi (128x128) ve baş pozisyonunu temsil eden 10 adet koordinat bilgisi girdi olarak verilmiştir. İlk üç girdi bakış yönünü temsil edecek uzamsal özelliklerin elde edilmesi için 4 evrişim katmanından

geçirilip tam bağlantı katmanına iletmek için düzeltilmiştir. Diğer iki girdide sırasıyla yüzün kameraya göre konumu, uzaklığı ve baş pozisyonunu temsil etmek üzere düzleştirip tam bağlantı katmanına iletilmiştir. Tam bağlantı katmanları hariç her katmanda toplu normalleştirme (batch normalization) ve maksimum havuzlama (max pooling) uygulanmıştır. Toplu normalleştirme, derin sinir ağlarının eğitimi sırasında değişen parametreler sonucunda katman girdilerinin dağılımlarının normalize edilmesi anlamına gelmektedir. Böylelikle daha yüksek öğrenme hızlarına ulaşılmaktadır [147]. Havuzlama, bir önceki katman çıktılarının uzamsal çözünürlüğünü, uygulanan filtreye göre düşüren ve uzamsal alt örneklemesini veren bir yöntemdir [148]. Havuzlama, ağıdaki parametreleri azaltır, hesaplama verimliliğini artırır, aşırı uyumu düzenler ve ağı derinliği değişmez [149]. En popüler havuzlama yöntemi olan maksimum havuzlama, ilgili alandaki maksimum değeri alır ve havuzlama alanındaki diğer tüm birimleri yok sayarak alanın en belirgin özelliğini öne çıkarır [150]. Tüm katmanlarda tasarlanan sistemin daha iyi öğrenmesi için doğrusal olmayan çıktılar veren ReLU aktivasyon fonksiyonu kullanılmıştır [151]. Tasarlanan sistem Şekil 5.5'te gösterilmiştir.



Şekil 5.5: Tasarlanan ESA tabanlı sistem

Tasarlanan sistemin eğitimi için hazırlanan ve her eğitim parametresi için 23.040 adet kayıt içeren veri seti eğitim ve test olmak üzere sırasıyla %90 ve %10 olarak iki bölüme ayrılmıştır. Tasarlanan sistemin eğitiminde doğru tahminler elde edilebilmesi için seçilecek kayıp fonksiyonu önemli yer tutmaktadır. Tasarlanan sistemde kayıp fonksiyonu olarak iki nokta arasındaki mesafeyi ölçen öklit mesafesi (euclidean distance) kullanılmıştır. Öklit mesafesi görseli Şekil 5.6’da ve öklit mesafesini veren eşitlik ise Denklem 5.2’de verilmiştir. Şekil 5.6’daki “d” ifadesi tahmin edilen nokta ile gerçek nokta arasındaki mesafe anlamına gelmektedir. Bu mesafe bize hata değerini verir. Kullanıcı bakış hata değeri hesaplanırken kullanıcı ekranda koordinatı (x_1, y_1) bilinen bir noktaya bakar. Kameradan elde edilen yüz, Dlib kütüphanesi aracılığıyla yer işaretlerine bölünür. Bu yer işaretleri aracılığıyla dönüm noktalarından bir dizi parametre (Yüz, Sol göz, Sağ göz, Yüz konumu, Baş pozisyonu bilgisi) elde edilir. Bu parametreler önceden eğitilmiş ESA tabanlı sisteme verilir ve ekranda bakış noktası (x_2, y_2) koordinatları üretilir. Üretilen bakış noktası koordinatı ekranda işaretlenir. İşaretlenen (x_2, y_2) bakış noktası koordinatı ile koordinatı ekranda bilinen (x_1, y_1) noktası arasındaki mesafe “öklid mesafesi” ile hesaplanır. Elde edilen piksel cinsinden mesafe ile hata değeri elde edilmektedir. Bu işlem ekrandaki 2.304 noktaya uygulanarak ortalama bir hata elde edilmektedir. Test doğruluğu, bu ortalama hataya göre hesaplanır.

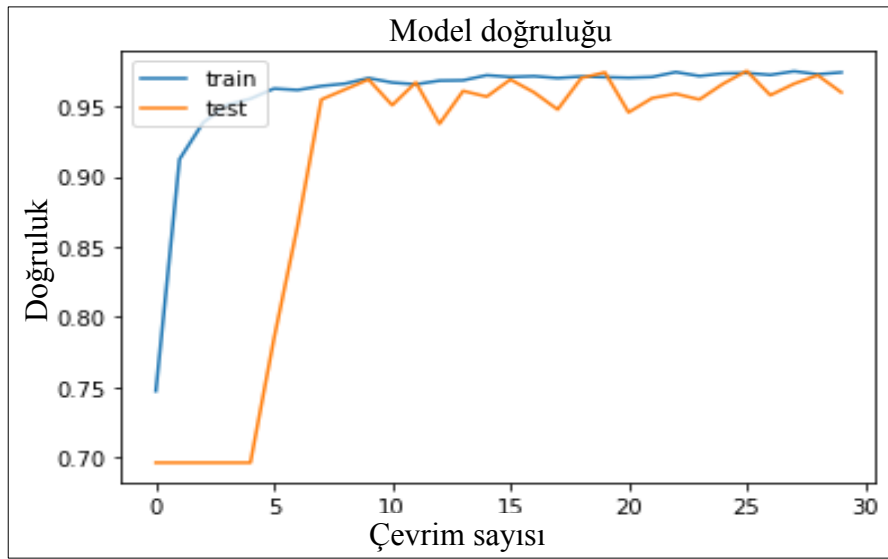


Şekil 5.6: Öklit mesafesi

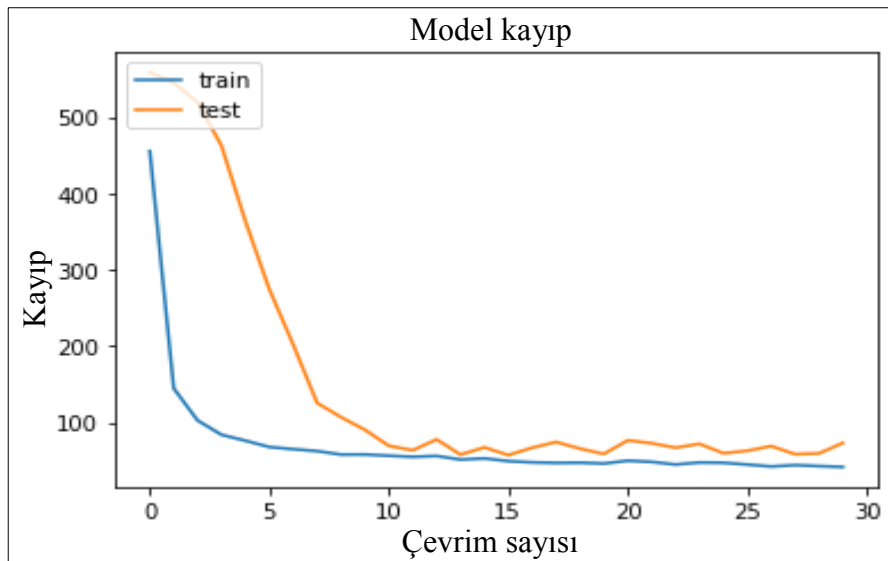
$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5.2)$$

Tasarlanan sistemin ESA modelinde optimizasyon algoritması olarak “Adam” kullanılmıştır [152]. Erken durdurmanın uygulandığı 29 çevrimde gerçekleşen eğitim sonunda ortalama 32 piksel eğitim ve ortalama 53 piksel test hatası elde edilmiştir.

Tasarlanan sistemin ESA modeli doğruluk ve kayıp fonksiyonlarına ait grafikler Şekil 5.7 ve Şekil 5.8’de gösterilmiştir. Modelin eğitim, test ve model parametre verileri de Çizelge 5.1’de verilmiştir. Şekil 5.7’deki grafik incelendiğinde eğitim ve test verileri sırasıyla %98-%96’lık bir doğruluk oranı elde edilmiştir. Bu doğruluk oranı ile tasarlanan sistemin doğru bir şekilde eğitildiği görülmektedir. Şekil 5.8’deki grafikte ise tasarlanan sistemin eğitimde kullanılan öklit mesafesi kayıp fonksiyonu ile hedeflenen bakış koordinatı ile sistem tarafından üretilen bakış koordinatı arasındaki farkın test aşamasında ortalama 53 piksel olduğu görülmektedir.



Şekil 5.7: Model doğruluk grafiği



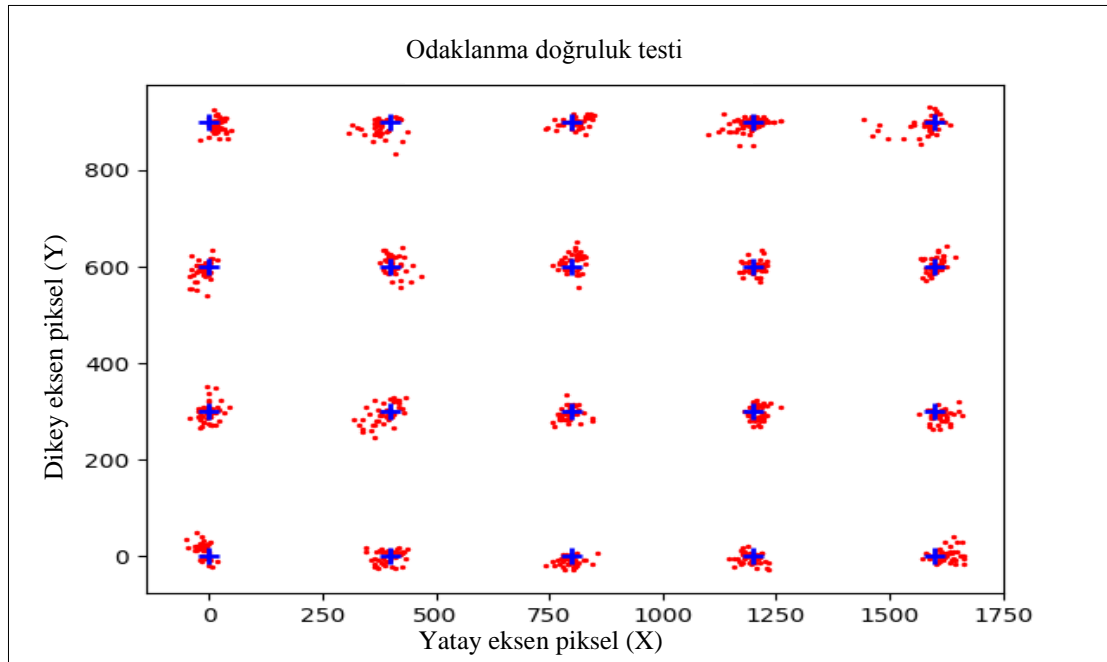
Şekil 5.8: Model kayıp grafiği

Çizelge 5.1: Sistem ESA modeli eğitim ve test verileri

Sistem ESA modeli eğitim ve test verileri	
Veriseti	23.040
Eğitim verisi	%90
Test verisi	%10
Giriş verisi	Yüz, sol ve sağ gözler, yüz konum maskesi, baş pozisyon bilgisi
Kayıp fonksiyonu	Öklit mesafesi
Optimizasyon yöntemi	Adam
Parti boyutu	64
Çevrim sayısı	29
Ortalama eğitim hatası	32 piksel
Ortalama test hatası	53 piksel

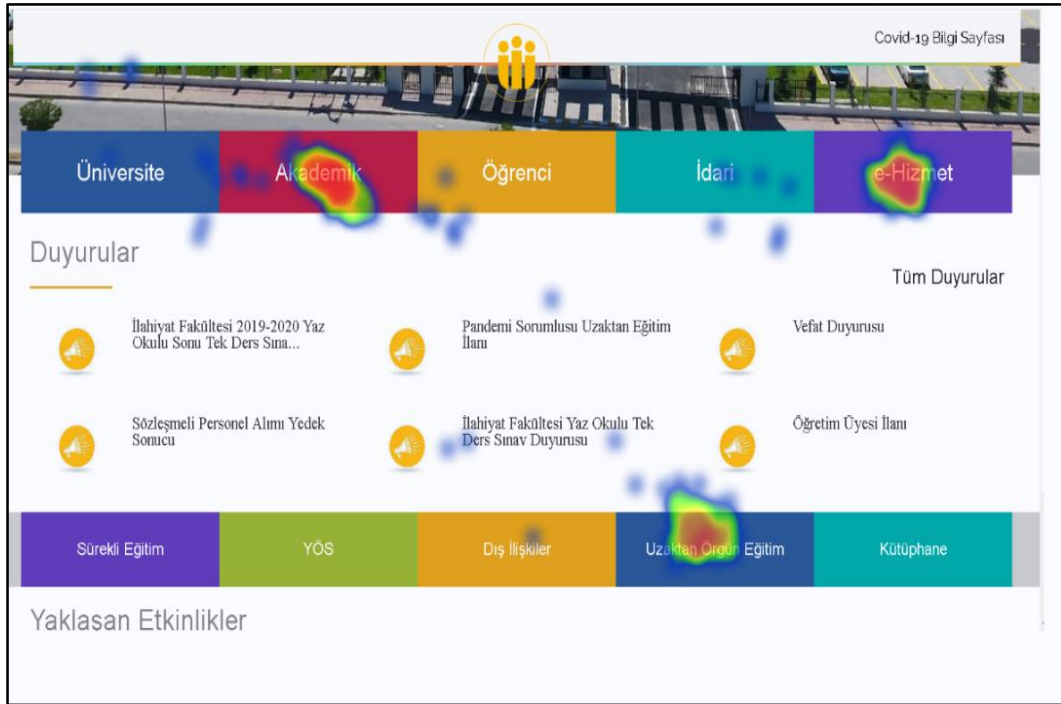
5.3 Deneysel Sonuçlar

Tasarlanan sistem, bakış takibinde eğitim aşamasında ortalama 32 piksel, test aşamasında ise ortalama 53 piksel hata ile sonuçlanmıştır. Tasarlanan sistemin uygulamada doğruluğunu görselleştirmek için ekrana yerleştirilen farklı koordinatlarda 20 farklı noktada odak testi yapılmıştır. Her noktanın 40 anlık görüntüsü dahil olmak üzere 800 anlık görüntü ile elde edilen bakış testi sonuçları Şekil 5.9'da gösterilmiştir.



Şekil 5.9: Odaklanma doğruluk testi

Eğitilen bakış takip sisteminin kullanılabilirliği bir web sayfası üzerinde bilgi arama deneyiyle test edilmiştir. Kullanıcıdan web sayfası üzerinde bulunan “Akademik”, “e-Hizmet” ve “Uzaktan Örgün Eğitim” sekmelerine sırasıyla 20’şer saniye bakması istenilmiştir. Deneysel sonunda kullanıcının göz hareketleri kaydedilerek web sayfası üzerinde baktığı yerler ısı haritası olarak görselleştirilmiştir. Isı haritaları, bir olgunun büyüklüğünü iki boyut üzerinde bir renk aralığı ile gösteren bir veri görselleştirme tekniğidir. Olgunun yoğunluğuna bağlı olarak renk değişir, bu da verinin nasıl değiştiği hakkında açık görsel ipuçları verir. Kırmızı en yoğun kümelemeyi mavi ise en düşük yoğunluğu gösterir [153]. Deneysel sonuçta kullanıcının daha çok sabitlendiği ve ilgilendiği yerler kırmızı olarak görselleştirilmiştir. Şekil 5.10’da kullanıcının bilgi arama deneyiyle bir web sayfası üzerindeki ilgilendiği alanların yoğunluğu ısı haritası ile gösterilmiştir.



Şekil 5.10: Bilgi arama deneyi ve ısı haritası gösterimi

Uygulamada veri seti hazırlanırken bazı sorunlar ile karşılaşılmıştır. Bu sorunlardan ilki python uygulamasındaki nesne hareketi esnasında nesnenin bulunduğu yere bakarken alınan anlık göz bakışının nesnenin bulunduğu konum ile aynı ekseninde olmadığı tespit edilmiştir. Bu durumundan kameranın aldığı anlık görüntü ile nesne hareketinin aynı zaman diliminde gerçekleşmediği anlaşılmaktadır. Bu durum eğitilecek olan modelin yanlış öğrenmesi anlamına gelmektedir. Bu sorunun üstesinden gelmek için her nesne hareketinde

kamera otomatik olarak kapatılıp yeniden aktif edilmiştir. Karşılaşılan diğer bir sorun ise uygulamada laptop kullanılmasıdır. Laptopa harici bir kamera yerleştirilmiştir. Dolayısıyla kamera ile laptop ekranının duruş açılarının tüm eğitim oturumlarında aynı olmasına gerektiği gözlemlenmiştir. Tasarlanan sistem gerçek dünya koşullarında her ne kadar başarılı sonuçlar verse de bazı kısıtlamalar hala mevcuttur. Bunlardan ilki eğitim verilerinin yüzün kameraya göre birçok farklı konumundan ve kameraya 40-70 cm uzaklıktan toplanmıştır. Kameranın sağ, sol, yukarı, aşağı yönünde 0-10 cm'lik baş hareket alanı sağlanmıştır. Sonuç olarak test uygulamasının da bu alanda yapılması daha iyi sonuçlar vermiştir. Eğitim verisinin kameraya göre daha geniş bir alanda ve farklı bakış açıları ile toplanmasıyla tasarlanan sistemin bakış tahmin doğruluğu baş hareketlerine daha az duyarlı olması sağlanacaktır. Bir diğer kısıtlama ise aydınlatma koşullarıdır. Uygulamadaki veri setinin günün farklı saatleri ve farklı aydınlatma koşulları altında toplanmasına dikkat edilmiştir. Bu şekilde modelin yüzde ve gözde oluşacak yansımalara karşıda daha az duyarlı olması sağlanmıştır. Yine de yüksek ve düşük aydınlatma koşulları sistemimizin çalışmasını olumsuz yönde etkilediği görülmüştür. Bir diğer kısıtlama ise veri setinin bir kullanıcıyla oluşturulmasıdır. Farklı kullanıcılar tarafından yapılan testlerde farklı hata oranları ile karşılaşılabilmektedir. Veri setine farklı kullanıcılarında verileri eklenerek sistemin daha işlevsel olması sağlanabilir. Ayrıca sistemin baş pozisyonlarına (pitch, yaw) karşı daha toleranslı olması için genel 3B yüz modelleme ve kamera yüz dönüm noktaları arasındaki ilişkiden baş pozisyonu hesaplanıp sisteme girdi olarak dahil edilmesiyle daha iyi sonuçlar elde edilebilir. Tasarlanan sistem bu kısıtlamalara rağmen web sayfalarında kullanılabilir olduğu görülmüştür.

6. BAKIŞ TAKİBİNİN ÇOCUKLARDA DİKKAT EKSİKLİĞİ TANISINDA KULLANIMI

Çocuklarda Dikkat Eksikliği ve Hiperaktivite Bozukluğu (DEHB) çocukların yaşam kalitelerini ve öğrenme süreçlerini olumsuz yönde etkilemektedir. Bu olumsuz etkinin araştırılması ve DEHB tanısının tespit edilmesi uzmanlar tarafından sağlanmaktadır. Ancak, uzmanların tanı koyma süreci hem geleneksel yöntemlere dayanmakta hem de uzmanın kişisel yorumuna göre değişkenlik gösterebilmektedir. Bu nedenle, DEHB tanısının konulmasında uzmanın kişisel görüşü ve yorumundan bağımsız olarak objektif testlere ihtiyaç duyulmaktadır. Tezin bu bölümünde, bireylerde DEHB tanısının objektif bir şekilde ortaya konulmasında bakış takibinin etkisi incelemiştir. Yapılan çalışmada (6-16) yaş arası 31 adet DEHB tanısı konulan çocuk ile 25 adet sağlıklı çocuktan oluşan kontrol grubuna bakış takip cihazı eşliğinde Sürekli Performans Testi (SPT) uygulanmıştır. Test sonunda tüm bireylere ait (hasta / kontrol grubu) göz hareket verileri kayıt altına alınmıştır. Kaydedilen veriler makine öğrenmesi yöntemiyle sınıflandırma analizine tabii tutulmuştur. Analiz sonunda elde edilen sonuçlar, bakış takibinin uzmanlara DEHB tanısı koymada yardımcı olabilecek bir karar destek sistemi olabileceğini ortaya koymuştur.

Bu bölümde ilk olarak bakış takibinin DEHB tanısında kullanımı üzerine literatür çalışmalarına kısaca yer verilmiştir. Daha sonra deneye ait katılımcı bilgileri, kullanılacak bakış takip cihazının teknik özellikleri, deney tasarımı ve ölçüm metrikleri detaylı bir şekilde açıklanmıştır. Sonrasında makine öğrenmesi tabanlı sistemin eğitilmesi için gerekli veri setinin oluşturulma adımları anlatılmıştır. Son olarak eğitilen modelin deneysel sonuçlarına yer verilmiştir.

6.1 İlgili Çalışmalar

Literatürde bakış takibinin DEHB’u olan hastaların tanısında kullanıldığı araştırmalardan bazıları şunlardır. Stokes ve diğ. [13] sanal gerçeklik teknolojisini kullandıkları çalışmada 8-12 yaşları arasında DEHB tanısı konulmuş 20 bireyle sanal sınıf ortamında dikkat dağınıklığı testi gerçekleştirmişlerdir. bakış takibi eşliğinde yapılan testte bireylerin dikkat dağınıklığının zamansal dinamikleri değerlendirilmiştir. Kullandıkları

yöntemde sanal sınıf ortamındaki beyaz tahta üzerinde verilen görevlere odaklanmaları ve mümkün olduğu kadar çok problemin doğru bir şekilde tamamlamaları talimatı verilmiştir. Bu görevler esnasında bireylerin dikkatlerini çekecek üç farklı çeldirici grubu kullanılmıştır. Görev sonunda yapılan değerlendirmelerde bireylerin çeldiricilere uzun süre odaklanmadığı fakat çeldiricilere odaklanıldığında beyaz tahta görevlerine dönmelerinde zorlandıklarını ve görev performanslarının düştüğünü gözlemlemişlerdir. De silva ve diğ. [154] göz hareket verileri kullanılarak kural tabanlı bir karar destek sistemi ile DEHB tanısının konulması üzerine bir deney çalışması yapmışlardır. Deneye, yaşları 18-65 arasında değişen DEHB tanısı konmuş 14 birey katılmıştır. Katılımcılara bakış takip cihazının yerleştirildiği bir ekranda okumaları için cümleler gösterilmiştir. Bu görev sürecinde katılımcıların sabitlenme/sakkad süreleri, sabitlenme/sakkad ortalamaları, sabitlenme/sakkad standart sapmaları ve sol/sağ göz bebeği çapları kayıt altına alınmıştır. Bu veriler kullanılarak karar destek sistemi ile DEHB tanısında %89 oranında bir kesinlik elde etmişlerdir. Elbaum ve diğ. [155] DEHB tanısında bireylerde bakış takibinin klinik karar verme sürecini geliştirebileceğini öne sürmüşlerdir. Bu amaçla MOXO-dCPT sürekli performans testi ile bir bakış takip cihazının beraber kullanıldığı bir çalışma yapmışlardır. Çalışma kapsamında 43 adet DEHB tanısı konmuş birey ile 42 adet sağlıklı birey performans testiyle karşılaştırılmıştır. Karşılaştırmalar sonunda DEHB’u olan bireylerin sağlıklı bireylere oranla ilgisiz alanlara daha fazla zaman harcadıkları görülmüştür. Testteki görsel çeldiricilerin DEHB’u olan bireylerin göz hareketlerini önemli ölçüde etkilediği görülmüştür.

6.2 Materyal ve Metodoloji

Yapılan çalışmada DEHB’li olan bireyler ile sağlıklı bireylere bilgisayar ekranı üzerinde belirli bir görev çerçevesinde animasyon gösterisi izletilmiştir. Animasyon gösterisi boyunca göz hareketlerinin takibi ve kayıt altına alınması için göz bebeği merkezi ve kornea yansıması çalışma prensibine dayanan Gazepoint GP3 HD bakış takip cihazı kullanılmıştır. Animasyon sonunda tüm bireylere ait göz hareket verilerinin DEHB tanısında ayırt edici özellikler olup olmadığı incelenmiştir. Son olarak özellik olarak kabul edilen ölçüm metrikleri DEHB’li bireylerin sınıflandırılması için bir Destek Vektör Makine’sinin eğitiminde kullanılmıştır.

6.2.1 Katılımcılar

Çalışmanın Akademik ve Etik Kurul raporları İnönü Üniversitesi Çocuk Psikiyatrisi Anabilim Dalı Başkanlığı ve İnönü Üniversitesi Fen ve Mühendislik Bilimleri Bilimsel

Araştırma ve Yayın Etik Kurulundan alınmıştır. Alanında uzman bir psikiyatrist tarafından DEHB tanısı konmuş 31 çocuk ile 25 adet kontrol grubu (sağlıklı çocuk) olmak üzere toplam 56 adet çocuk üzerinde çalışma gerçekleştirilmiştir. Çalışma için seçilen en küçük çocuğun yaşı 6, en büyük çocuğun yaşı ise 16'dır. Kontrol grubunda bulunan çocukların hiç hiçbirinde daha önceden bildirilen bir DEHB öyküsü saptanmamıştır. Tüm katılımcılar, deney öncesinde göz izleme tekniğini kullanan herhangi bir teste katılmamıştır.

6.2.2 Cihazlar ve teknik özellikleri

Çalışmanın dikkat testinde, uyarıların sunulacağı 1600×900 çözünürlüğe sahip bir 22" boyutunda ekran, Intel Core i7 işlemci, NVIDIA GeForce 405 4GB ekran kartı, 8 GB Ram ve Windows 10 işletim sistemine sahip bilgisayar kullanılmıştır. Ayrıca deney esnasında katılımcıların göz hareketlerinin kayıt altına alınmasında yüksek doğrulukta 150 Hz (saniye 150 göz hareket verisi) ile çalışan Gazepoint firmasının ticari bakış takip cihazı Gazepoint GP3 HD kullanılmıştır. Bu cihaz Tıbbi araştırma ve teşhis, kurumsal araştırma ve geliştirme, akademik araştırmalar ve nöropazarlama gibi birçok farklı alanda çözümler sunmaktadır. Çizelge 6.1 ve Şekil 6.1'de sırasıyla çalışmada kullanılan Gazepoint firmasına ait GP3 HD bakış takip cihazının teknik bilgileri ve cihaza ait görsel verilmiştir.

Çizelge 6.1: GP3 HD bakış takip cihazı teknik özellikleri

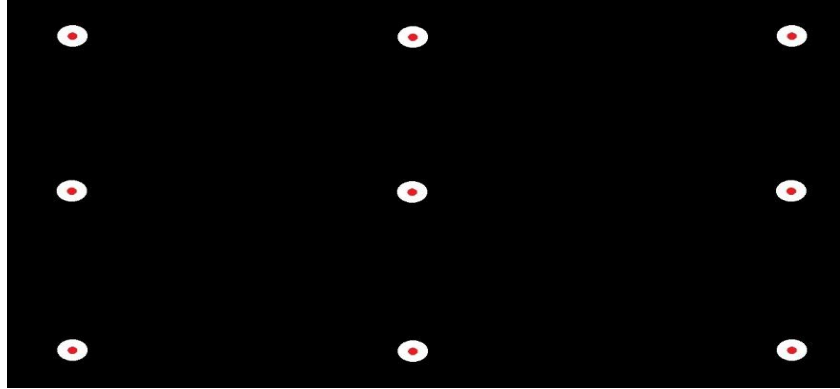
Özellik	Açıklama
Görsel açı doğruluğu	0.5-1.0 derece
Örnekleme hızı	60Hz-150Hz
Kalibrasyon	9 nokta
Baş hareketi	35cm (yatay) x 22cm(dikey) hareket
Baş derinlik hareketi	±15 cm
Ekran	24" veya daha küçük ekranlarla uyumlu
Sistem gereksinimleri	Intel Core i5 – 8. nesil veya daha hızlı, 8 GB RAM, Windows 7, 8.1 veya 10, USB3.0 bağlantı noktası



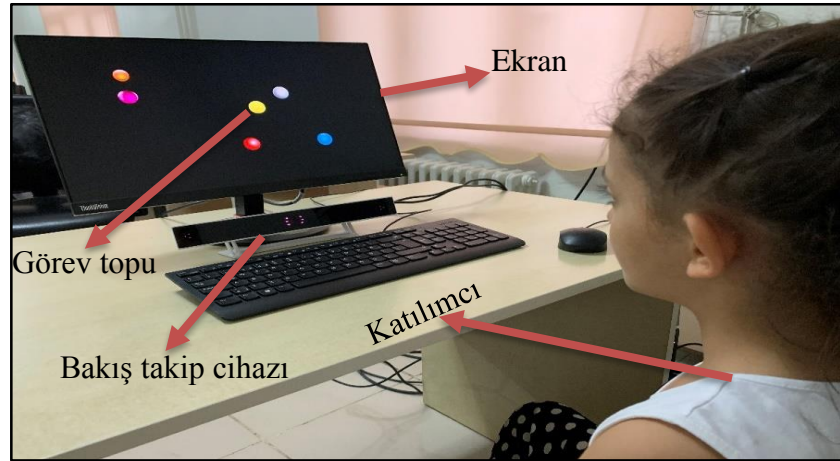
Şekil 6.1: GP3 HD bakış takip cihazı

6.2.3 Deney tasarımı ve ölçüm metrikleri

Bakış takip cihazı, deneye başlamadan önce deneyin gerçekleştirileceği ekranın alt kısmına yerleştirilmiştir. Katılımcılar, ekranın yaklaşık 55-65 cm önünde deneye alınmıştır. GP3 HD cihazının 9 noktalı kalibrasyon prosedürü kullanılarak her katılımcı için kalibre edilmiştir. Katılımcılardan deneye alınacakları ekran üzerinde 9 farklı noktada sırasıyla görünen kırmızı noktalara odaklanmaları istenmiştir. Bu prosedür ile bakış takibi daha doğru bir şekilde gerçekleştirilmektedir. Katılımcılardan deney esnasında başlarını hareket ettirmemeleri ve ekran dışına bakmamaları için hatırlatmalar yapıldı. Deney yazılımı, çocuklar tarafından ilgi çekici olması sebebiyle Python yazılım geliştirme ortamında oluşturulan bir animasyon gösterisinden oluşmaktadır. Böylelikle animasyonda verilecek görevlere daha uzun süre devam etmesine ortam sağlanmıştır. Kalibrasyon ekranı ve katılımcının deney esnasındaki oturma ve cihaz düzeni sırasıyla Şekil 6.2 ve Şekil 6.3'te gösterilmiştir.

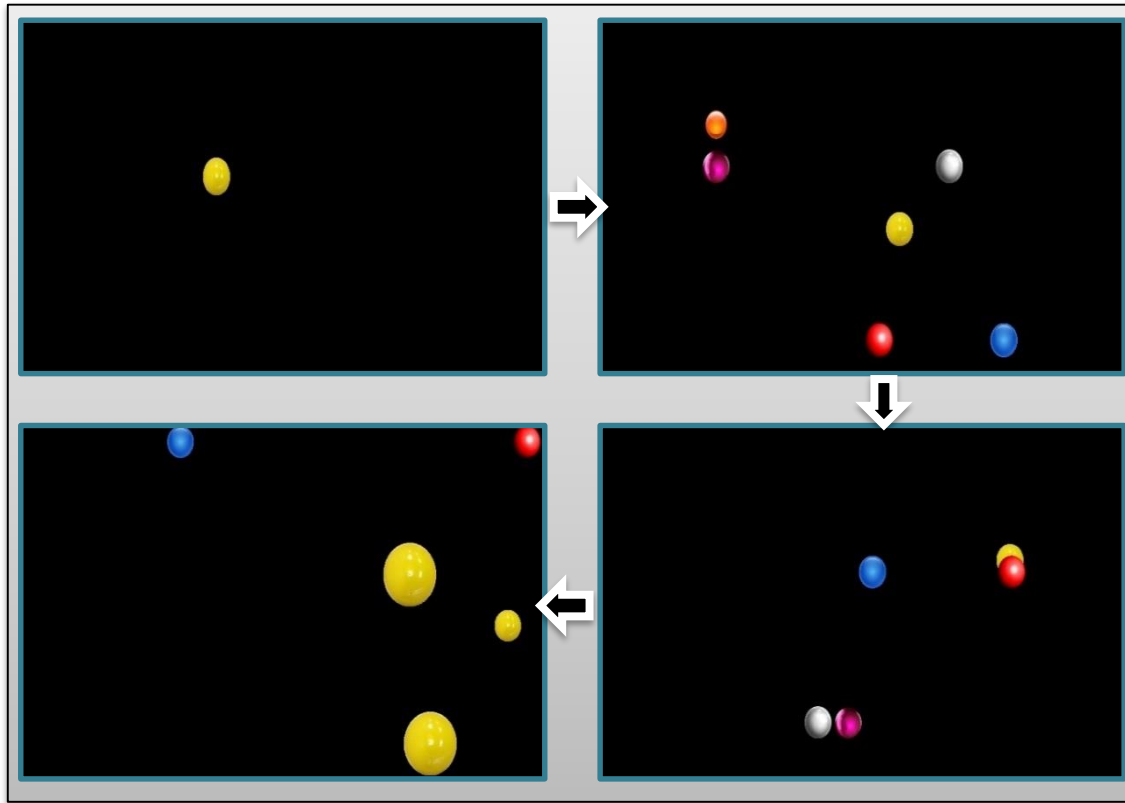


Şekil 6.2: Kalibrasyon ekranı



Şekil 6.3: Katılımcı ve cihaz düzeni

Animasyonun başında katılımcıya ait cinsiyet, yaş ve DEHB tanısı bilgileri alınmaktadır. Animasyon 74 saniye süren top hareketlerinden oluşmaktadır. Katılımcılardan animasyonun başında sahnede beliren sarı topun (görev topu) animasyon süresi boyunca takip edilmesi görevi istenmektedir. Bu süre boyunca görev topu dışında belirli bir süre sonunda hareket eden başka toplar belirecektir. Hareket eden topların renkleri, büyüklüğü ve yönleri katılımcının dikkatinin dağıtılması amacıyla senaryo çerçevesinde çeldiriciler olarak kullanılmıştır. Bu senaryoda genel olarak şunlara dikkat edilmiştir. Toplar sahnenin farklı bölgelerinden farklı renklerde, farklı yönlerde ve farklı hızlarda hareket etmeleri, topların bir kısmı görev topuna uzakta hareket ederken bir kısım top ise görev topuna yakın hatta çakışacak şekilde hareket etmesi sağlanmıştır. Görevin sonlarına doğru görev topu ile aynı renkte ve görev topundan daha büyük çeldirici toplar sahneye dahil edilmiştir. Bu senaryoya ait görseller Şekil 6.4'te gösterilmiştir.



Şekil 6.4: Görev topu senaryosu

Katılımcıların görev süresi boyunca gerçekleştirdiği göz hareketleri, bakış takip cihazıyla kayıt altına alınmıştır. Görev esnasında gerçekleştirilen göz hareketlerinin takibi, katılımcının ekran üzerindeki göz sabitlemeleri (fiksasyon-fixation) metriği dikkate

alınarak ölçülmüştür. Ekran üzerindeki göz sabitlemeleri katılımcının ilgi alanına işaret etmektedir. Çalışmada 4 adet ölçüm metriği DEHB tanısı için ayırt edici özellikler olarak seçilmiştir. Bu metrikler sırasıyla katılımcıların yaşı, katılımcının görev topu üzerindeki sabitlemelerden elde edilen İlgi Alanı Süresi (İA-Dwell time), İlgi Alanı Dışı Süre (İAD), görev topuna Tekrar Ziyaret Sayısı (TZ-Revisit) olarak belirlenmiştir. Çizelge 6.2’de deney sonrası hasta (H) ve sağlıklı (S) katılımcılardan elde edilen bakış takip cihazına ait ölçüm metriklerinden bir kısım verilmiştir.

Çizelge 6.2: Katılımcılara ait ölçüm metrik değerleri

Id	Cinsiyet	Yaş	İA	İAD	TZ	DEHB
1	E	9	66	8	7	H
2	E	11	58	16	9	H
3	E	11	44	30	31	H
4	E	11	48	26	24	H
5	K	12	65	9	3	H
6	K	10	39	35	12	H
7	E	9	47	27	16	H
8	E	7	54	20	19	H
9	E	15	30	44	12	H
10	K	8	58	16	5	S
11	K	13	64	10	5	S
12	K	8	53	21	7	S
13	E	11	59	15	6	S
14	E	6	61	13	11	S
15	E	11	59	15	6	S
16	K	12	62	12	4	S
17	K	12	65	9	3	S
18	E	16	69	5	4	S
.
.
.
56	K	12	60	14	4	S

Deney veri seti incelendiğinde katılımcılara ait 5 adet özellik bulunmaktadır. Bu özelliklerden cinsiyet özelliği 38 (erkek), 18 (kız) olacak şekilde dengeli bir dağılım olmadığından özellik olarak alınmamıştır. Bunu yanında İA, İAD ve TZ özellikleri incelendiğinde Çizelge 6.3’te görüldüğü gibi hasta ve sağlıklı ortalamaları bu özelliklerin ayırt edici olabileceği yönünde fikir vermektedir.

Çizelge 6.3: Hasta ve sağlıklılara ait ölçüm ortalamaları

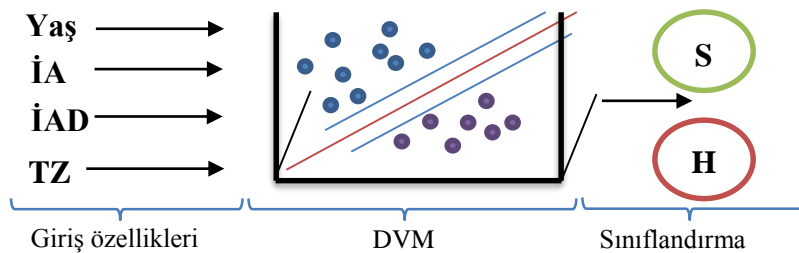
Tanı	İA ortalama (sn)	İAD ortalama (sn)	TZ ortalama (adet)
Hasta	38.23	35.77	12.12
Sağlıklı	62.48	11.52	3.56

Çizelge 6.3 incelendiğinde hasta bireylere ait görev alanına odaklanma süresinin ortalaması sağlıklı bireylere göre daha düşük olduğu görülmüştür. Bu durumun dikkat dağıtıcı olarak senaryoya dahil edilen çeldiricilere sağlıklı bireylere oranla daha fazla odaklanıldığını göstermektedir.

6.2.4 Destek Vektör Makinesi ve eğitimi

DVM [156], sınıflandırma ve regresyon görevleri için kullanılan popüler bir denetimli makine öğrenimi algoritmasıdır. Öncelikle ikili sınıflandırma problemlerini çözmek için kullanılır, ancak çok sınıflı sınıflandırmayı ele alacak şekilde genişletilebilir. DVM'ler, hem doğrusal olarak ayrılabilen hem de doğrusal olarak ayrılamayan verilerle başa çıkmada etkilidir. Bir DVM'nin temel amacı, özellik uzayındaki farklı sınıflara ait veri noktalarını ayıran optimal bir hiper düzlem bulmaktır.

Tüm bireylerden elde edilen ölçüm verileri DVM'ye verilmek üzere normalizasyon işlemine tabii tutulmuştur. Böylelikle farklı ölçeklere sahip verilerin belirli bir aralığa veya dağılıma dönüştürülmesiyle veriler arasındaki ölçek farklılıklarının ortadan kaldırılması veya azaltılması amaçlanmıştır. Bireylere ait yaş, İA, İAD ve TZ verileri DEHB tanısı için sınıflandırılmak üzere Destek Vektör Makinesine verilmiştir. DVM'nin bu özelliklere göre eğitim ve DEHB sınıflandırmasını yapabilmesi için ilişkilendirilmesi gereken hedef etiketler hasta bireyler için "0", sağlıklı bireyler için "1" olarak alınmıştır. Bireylerden elde edilen veri setinin büyüklüğü DVM'nin daha doğru sonuçlar üretmesinde önemli bir etkidir. Dolayısıyla 56 bireyden elde edilen ölçüm metriklerinin DVM ile eğitiminde daha objektif sonuçlar vermesi için veri setimize, makine öğrenimi modellerinin performansını değerlendirmek ve genellebilirliklerini test etmek için kullanılan bir yöntem olan çapraz doğrulama (cross-validation) uygulanmıştır. Şekil 6.5'te DVM tabanlı sisteme ait görsel verilmiştir.



Şekil 6.5: DVM tabanlı sistem

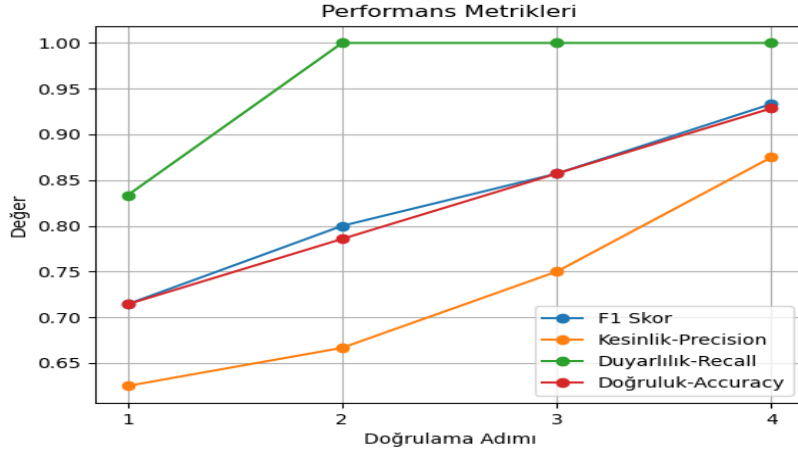
6.3 Deneysel Sonular

DVM nin performansını deęerlendirmek iin 4 kat apraz doęrulama (4-fold cross-validation) yntemi uygulanmıřtır. Bu yntem ile, mevcut veri setini drt eřit paraya bler ve her bir parayı sırayla test veri seti olarak kullanırken dięer  parayı eęitim veri seti olarak kullanır. Veri seti blme iřleminde veri seti rastgele olarak drt eřit paraya blnr. İlk iterasyonda, bir para test veri seti olarak seilir ve dięer  para eęitim veri seti olarak birleřtirilir. Model, eęitim veri seti zerinde eęitilir ve test veri seti zerinde deęerlendirilir. Seilen performans metriklerine (doęruluk, kesinlik, duyarlılık) dayanarak modelin performansı llr. İkinci iterasyonda, test ve eęitim veri seti paraları deęiřir. İřlem 4 kez tekrarlanır, bylece her bir para test verisi olarak kullanılır. Sonu olarak, her bir iterasyon sonucunda elde edilen performans lmleri bir araya getirilir ve ortalama performans deęeri hesaplanır. Bu, modelin genel performansını daha gvenilir bir řekilde deęerlendirmemize yardımcı olur.

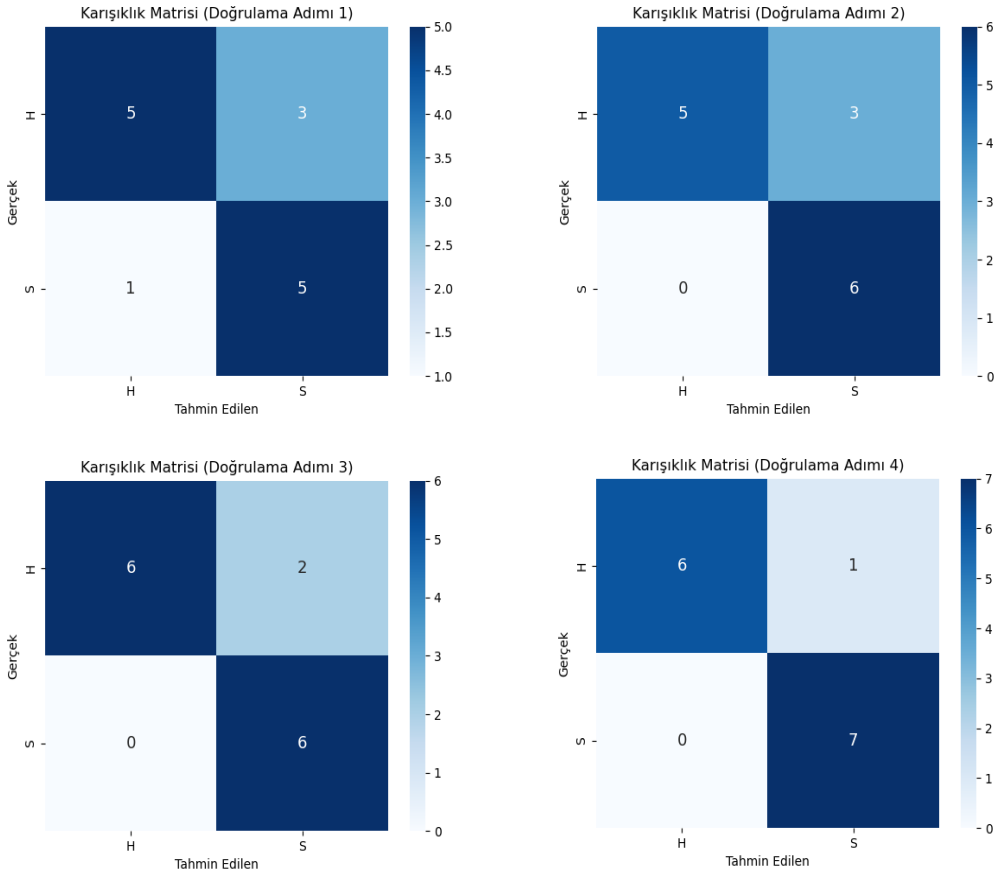
Performans metriklerinden Doęruluk (accuracy), bir sınıflandırma modelinin doęru tahminlerin oranını len bir performans metrięidir. Doęruluk, toplam doęru tahmin sayısının toplam veri noktası sayısına blnmesiyle hesaplanır. Yani, doęru tahminlerin (true positives ve true negatives) toplam veri noktalarına oranını gsterir. Doęruluk, genel model performansını deęerlendirmek iin kullanılır. Ancak, dengesiz sınıf daęılımları durumunda yanıltıcı olabilir. Kesinlik (precision) ise bir sınıflandırma modelinin pozitif olarak tahminledięi rneklerin ne kadarının gerekten pozitif olduęunu ler. Kesinlik, gerek pozitiflerin (true positives) toplam tahmin edilen pozitifler (true positives ve false positives) zerine blnmesiyle hesaplanır. Kesinlik, yanlış pozitif tahminlere (false positives) odaklanır. Duyarlılık (recall), gerek pozitiflerin ne kadarının doęru bir řekilde saptandıęını len bir performans metrięidir. Duyarlılık, gerek pozitiflerin (true positives) toplam gerek pozitifler (true positives ve false negatives) zerine blnmesiyle hesaplanır. Duyarlılık, yanlış negatiflere (false negatives) odaklanır ve modelin ne kadar iyi bir řekilde pozitif sınıfı tanıdıęını deęerlendirir. F1 skoru, bir sınıflandırma modelinin kesinlik (precision) ve duyarlılık (recall) metriklerini birleřtirerek hesaplanan bir performans lsdr. F1 skoru, bu iki metrięi dengeleyerek eřit dengeli daęılmayan veri kmlerinde modelin performansını daha kapsamlı bir řekilde deęerlendirmeye yardımcı olur. F1 skor eřitlięi Denklem 6.1’de verilmiřtir.

$$F1 = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (6.1)$$

DVM modelinin veri seti ile 4 kat çapraz doğrulama sonucunda ortalama %82.61'lik F1 skor sınıflandırma doğruluğuna ulaştığı tespit edilmiştir. Çapraz doğrulama adımlarına ait performans metrikleri ve karışıklık matrisleri sırasıyla Şekil 6.6 ve Şekil 6.7'de verilmiştir.



Şekil 6.6: Performans metrikleri



Şekil 6.7: Karışıklık matrisleri

Karışıklık matrisleri incelendiğinde, DVM modelinde çapraz doğrulamanın her adımı için 14 adet veri test işlemine tabi tutulmuştur. Doğrulamanın 4 adımında da sağlıklı bireylerin yüksek doğrulukta sınıflandırıldığı görülmektedir. Sadece 1. doğrulama adımında 1 adet sağlıklı birey DEHB'li olarak tahmin edilmiştir. Fakat modelin DEHB'li bireylerin sınıflandırılmasında doğruluk oranı sağlıklı bireyelere oranla nispeten daha düşüktür. Modelin 1. ve 2. doğrulama adımlarında sırasıyla 3'er DEHB'li birey "Sağlıklı" olarak tahmin edilirken, 3. ve 4. doğrulama adımlarında sırasıyla 1 ve 2 DEHB'li birey "Sağlıklı" olarak tahmin edilmiştir. Tüm doğrulama adımları göz önüne alındığında ise DEHB'li ve Sağlıklı bireyler ortalama %82.61 oranında doğru bir şekilde tahmin edilmiştir.

Bu çalışmada bakış takip cihazı eşliğinde gerçekleştirilen sürekli performans testinde katılımcılara verilen görev topu takibinde elde edilen göz hareketlerine ait ölçüm metriklerinin dikkat eksikliği tanısı üzerindeki belirleyiciliği DVM algoritması ile analiz edilmiştir. Katılımcılara ait göz hareket verilerinin DVM algoritmasına, DEHB tanısında ayırt edici olmasını sağlamak için test esnasında görev topunun yanında dikkat dağıtıcı unsurlar eklenmiştir. Deneyde katılımcıların yaşı, görev topuna odaklanma süresi, görev topu dışında geçirilen süre ve görev topuna tekrar ziyaret ölçüm metrikleri DVM sınıflandırma modeli için özellik olarak alınmıştır. Katılımcılardan elde edilen özellikler normalizasyon işleminden sonra DVM modelinde 4 kat çapraz doğrulama ile eğitilmiş ve ortalama sınıflandırma F1 skor doğruluğu elde edilmiştir. Sınıflandırma sonunda elde edilen sonuçlar mevcut veri seti ve özellikler ile bakış takip verilerinin DEHB tanısında %82.61 oranında belirleyici olduğu görülmüştür. Bu bilgiler ışığında DEHB'li bireylerin, kontrol grubu bireylere göre ekran üzerinde dikkat dağıtıcı nesnelere bakarak önemli ölçüde zaman harcadığı gözlemlenmiştir. Elde edilen başarının çalışmada kullanılan veri sayısı ve dikkate alınan özellikler ile sınırlı olduğuna dikkat edilmelidir. DVM modelinin daha iyi bir genelleme yeteneği sergilemesi için katılımcı sayısının artırılması ve bunun yanında ilk fiksasyon süresi, fiksasyon süresi, ortalama fiksasyon süresi, fiksasyon sayısı, sakkad genliği, sakkad sayısı gibi farklı ölçüm metriklerinin sürece dahil edilmesi öngörülmektedir.

7. SONUÇLAR

Bakış takibi, gelişen teknolojiyle beraber insan-bilgisayar etkileşimi alanında yaygın olarak kullanılmaya başlayan önemli teknolojilerden biridir. Bu teknoloji birçok farklı alandaki problemlere çözümler sunmaktadır. Günümüzde bu teknolojiyi ticari bir ürüne dönüştüren firmalar bulunmaktadır. Bakış takibinin kullanıldığı araştırma-geliştirme ve akademik çalışmalarda genellikle bu firmalara ait yüksek bakış takip doğruluğu sunan cihazlar kullanılmaktadır. Söz konusu cihazlar yüksek çözünürlük ve örnekleme hızlarına sahip kameralar ve kızılötesi ışık kaynağından oluşmaktadır. Bu cihazlar yüksek donanım gereksinimlerinden dolayı maliyetlidir. Yüksek maliyetinden dolayı bu cihazların kullanım alanı araştırma geliştirme ve akademik çalışma alanlarıyla sınırlıdır. Yüksek doğrulukta çalışan bu cihazların kullanımı esnasında bazı kısıtlamalar mevcuttur. Bu kısıtlamalar, her kullanıcı için kalibrasyon prosedürünün uygulanması, ekran ve kullanıcı arası belirli mesafenin olması ve uygun aydınlatma koşullarının mevcut olması olarak özetlenebilir.

Bakış takip teknolojisinin herkes tarafından deneyimlenebilmesi için cep telefonu, tablet ve kişisel bilgisayarlarda tüketici sınıfı kameralar ile çalışabilen alternatif yöntemlere ihtiyaç vardır. Bu amaçla bu tez kapsamında bakış takip teknolojisinin herkes tarafından kullanılabilmesi için derin öğrenme tabanlı bir sistem tasarlanmıştır. Bu sistemin ticari bakış takip cihazlarının aksine düşük maliyetli olması, kalibrasyon gerektirmemesi ve farklı aydınlatma koşullarına karşı sağlam bir performans sergilemesi açısından geliştirilen sistem tezin en önemli yönünü oluşturmaktadır. Sistemin kalibrasyon gerektirmemesi ve farklı ortam koşullarına karşı sağlam performans sergilemesi derin öğrenme ağlarının örüntü tanıma gücüne dayanmaktadır. Tez kapsamında derin öğrenme tabanlı bakış takip sisteminin dışında bakış takip sistemlerinin önemli parametrelerinden olan göz tespiti ve göz bebeği merkezi lokalizasyonu için de derin öğrenme tabanlı yöntemler geliştirilmiştir. Ayrıca derin öğrenme tabanlı bakış takip sistemlerinin sağlamlığının ve doğruluğunun daha iyi anlaşılması için tez kapsamında görüntü işleme tekniklerine dayanan farklı bir bakış takip sistemi de geliştirilmiştir. Tezin son bölümünde bakış takip sistemi verileri ile bireylerde DEHB tanısı koymada yardımcı bir karar destek sistemi olabileceğini gösteren Destek Vektör Algoritması'na dayanan bir yöntem geliştirilmiştir.

KAYNAKLAR

- [1] **Vanichvasin, P.** (2020). Effects of Visual Communication on Memory Enhancement of Thai Undergraduate Students, Kasetsart University, *Higher Education Studies*, 11(1), 34.
- [2] **Wade, N.J.** (2010). Pioneers of eye movement research, *I-Perception*. 1(2), 33.
- [3] **Market Research Reports & Consulting | Grand View Research, Inc.** Erişim: 02 Ağustos 2022, <https://www.grandviewresearch.com/>
- [4] **Facebook Purchases An Eye-Tracking Startup With An Eye Toward Enhancing VR.** Erişim: 14 Ağustos 2022, <https://www.forbes.com/sites/kevinmurnane/2016/12/29/facebook-purchases-an-eye-tracking-startup-with-an-eye-toward-enhancing-vr/?sh=5c8560bd544b>
- [5] **Carelli, L., Solca, F., Tagini, S., Torre, S., Verde, F., Ticozzi, N., et al.** (2022). Gaze-Contingent Eye-Tracking Training in Brain Disorders: A Systematic Review, *Brain Sciences*, 12(7), 931.
- [6] **Jarodzka, H., Skuballa, I. & Gruber, H.** (2021). Eye-Tracking in Educational Practice: Investigating Visual Perception Underlying Teaching and Learning in the Classroom, *Educational Psychology Review*, 33(1), 1–10.
- [7] **Ichindelean, M.T., Teodora, M., Cetină, I., Orzan, G., Sánchez-Fernández, J., Casado-Aranda, L.-A., et al.** (2021). A Comparative Eye Tracking Study of Usability—Towards Sustainable Web Design, *Sustainability* 2021, 13(18), 10415.
- [8] **Higgins, E., Leinenger, M. & Rayner, K.** (2014). Eye movements when viewing advertisements, *Frontiers in Psychology*, 5, 210.
- [9] **Białowas, S. & Szyszka, A.** (2019). Eye-tracking in Marketing Research, *Managing Economic Innovations – Methods and Instruments*, 91–104.
- [10] **Antunes, J. & Santana, P.** (2018). A study on the use of eye tracking to adapt gameplay and procedural content generation in first-person shooter games, *Multimodal Technologies and Interaction*, 2(2).
- [11] **Ojsteršek, T.C. & Topolšek, D.** (2019). Eye tracking use in researching driver distraction: A scientometric and qualitative literature review approach, *Journal of Eye Movement Research*, 12(3), 1–30.
- [12] **Paing, M.P., Juhong, A. & Pintavirooj, C.** (2022). Design and Development of an Assistive System Based on Eye Tracking, *Electronics* 2022, 11(4), 535.
- [13] **Stokes, J.D., Rizzo, A., Geng, J.J. & Schweitzer, J.B.** (2022). Measuring Attentional Distraction in Children With ADHD Using Virtual Reality Technology With Eye-Tracking, *Frontiers in Virtual Reality*. 3, 23.

- [14] **What is eye tracking? | How eye tracking works - Tobii.** Erişim: 14 Ağustos 2022, <https://www.tobii.com/group/about/this-is-eye-tracking/>
- [15] **Göz- Vikipedi.** Erişim: 12 Temmuz 2022, <https://tr.wikipedia.org/wiki/G%C3%B6z>
- [16] **Willoughby, C.E., Ponzin, D., Ferrari, S., Lobo, A., Landau, K. & Omidi, Y.** (2010). Anatomy and physiology of the human eye: effects of mucopolysaccharidoses disease on structure and function – a review, *Clinical & Experimental Ophthalmology*, 38(s1), 2–11.
- [17] **Human eye- Wikipedia.** Erişim: 30 Temmuz 2022, https://en.wikipedia.org/wiki/Human_eye
- [18] **Optical Axes and Angle Kappa- EyeWiki.** Erişim: 22 Eylül 2022, https://eyewiki.aao.org/Optical_Axes_and_Angle_Kappa
- [19] **Zhu, Z. & Ji, Q.** (2007). Novel eye gaze tracking techniques under natural head movement, *IEEE Transactions on Biomedical Engineering*, 54(12), 2246–2260.
- [20] **iMotions: Unpack Human Behavior.** Erişim: 3 Ağustos 2022, <https://imotions.com/>
- [21] **Rayner, K.** (2009). Eye movements and attention in reading, scene perception, and visual search, *Quarterly Journal of Experimental Psychology (2006)*, 62(8), 1457–1506.
- [22] **Learn about the different types of eye movement- Tobii Pro.** Erişim: 10 Ağustos 2022, <https://www.tobiipro.com>
- [23] **Land, M. & Tatler, B.** (2009). Looking and Acting: Vision and eye movements in natural behaviour, *Looking and Acting: Vision and Eye Movements in Natural Behaviour*, 1–288.
- [24] **Camcı, H.** (2017). Eye Tracking System and Orthodontics: Review, *Turkiye Klinikleri Journal of Dental Sciences*, 23(3), 184–190.
- [25] **Duchowski, A.T.** (2017). *Eye tracking methodology: Theory and practice* (3rd ed., pp.1–366).
- [26] **Robinson, D.A.** (1963). A Method of Measuring Eye Movement Using a Scleral Search Coil in a Magnetic Field, *IEEE Transactions on Bio-Medical Electronics*, 10(4), 137–145.
- [27] **Collewijn, H., van der Mark, F. & Jansen, T.C.** (1975). Precise recording of human eye movements, *Vision Research*, 15(3), 447-450.
- [28] **Whitmire, E., Trutoiu, L., Cavin, R., Perek, D., Scally, B., Phillips, J., et al.** (2016). EyeContact: Scleral coil eye tracking for virtual reality, *International Symposium on Wearable Computers, Digest of Papers*, (pp.184–191). September 12-16.

- [29] **Massin, L., Nourrit, V., Lahuec, C., Seguin, F., Adam, L., Daniel, E., et al.** (2020). Development of a new scleral contact lens with encapsulated photodetectors for eye tracking, *Optics Express*, 28(19), 28635.
- [30] **Guestrin, E.D. & Eizenman, M.** (2006). General theory of remote gaze estimation using the pupil center and corneal reflections, *IEEE Transactions on Biomedical Engineering*, 53(6), 1124–1133.
- [31] **Wang, J.G. & Sung, E.** (2002). Study on eye gaze estimation, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 32(3), 332–350.
- [32] **Ishikawa, T.** (2004). Passive driver gaze tracking with active appearance models.
- [33] **Matsumoto, Y. & Zelinsky, A.** (2000). An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement, *Proceedings - 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000*, (pp.499–504).
- [34] **Khan, M.Q. & Lee, S.** (2019). Gaze and Eye Tracking: Techniques and Applications in ADAS, *Sensors (Basel, Switzerland)*, 19(24).
- [35] **Majaranta, P. & Bulling, A.** (2014). Eye Tracking and Eye-Based Human–Computer Interaction, 39–65.
- [36] **Ye, L., Zhu, M., Xia, S. & Pan, H.** (2014). Cascaded convolutional neural network for eye detection under complex scenarios, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8833, 473–480.
- [37] **Murthy, L.R.D., Brahmhatt, S., Arjun, S. & Biswas, P.** (2021). I2DNet- Design and Real-Time Evaluation of Appearance-based gaze estimation system, *Journal of Eye Movement Research*, 14(4), 1–15.
- [38] **Morimoto, C.H., Amir, A. & Flickner, M.** (2002). Detecting eye position and gaze from a single camera and 2 light sources, *Proceedings - International Conference on Pattern Recognition*, 16(4), 314–317.
- [39] **Uhm, K.H., Kang, M.C., Kim, J.Y. & Ko, S.J.** (2020). Improving the robustness of gaze tracking under unconstrained illumination conditions, *Multimedia Tools and Applications*, 79(29–30), 20603–20616.
- [40] **Yoo, D.H. & Chung, M.J.** (2005). A novel non-intrusive eye gaze estimation using cross-ratio under large head motion, *Computer Vision and Image Understanding*, 98(1), 25–51.
- [41] **Gazepoint- Best Eye Tracking Software For Marketers.** Erişim: 10 Ağustos 2022, <https://www.gazept.com/>
- [42] **Wang, J.G., Sung, E. & Venkateswarlu, R.** (2005). Estimating the eye gaze from one eye, *Computer Vision and Image Understanding*, 98(1), 83–103.

- [43] **Chen, J. & Ji, Q.** (2008). 3D gaze estimation with a single camera without IR illumination, *Proceedings- International Conference on Pattern Recognition*.
- [44] **Reale, M.J., Canavan, S., Yin, L., Hu, K. & Hung, T.** (2011). A multi-gesture interaction system using a 3-D iris disk model for gaze estimation and an active appearance model for 3-D hand pointing, *IEEE Transactions on Multimedia*, 13(3), 474–486.
- [45] **Ronsse, R., White, O. & Lefèvre, P.** (2007). Computation of gaze orientation under unrestrained head movements, *Journal of Neuroscience Methods*, 159(1), 158–169.
- [46] **Pogalin, E., Redert, A., Patras, I. & Hendriks, E.A.** (2006). Gaze tracking by using factorized likelihoods particle filtering and stereo vision, *Proceedings - Third International Symposium on 3D Data Processing, Visualization, and Transmission, 3DPVT 2006*, (pp.57–64).
- [47] **Baluja, S. & Pomerleau, D.** (1993). Non-intrusive gaze tracking using artificial neural networks, in: *NIPS'93: Proceedings of the 6th International Conference on Neural Information Processing Systems*, (pp.753–760).
- [48] **Fischer, T., Chang, H.J. & Demiris, Y.** (2018). RT-GENE: Real-time eye gaze estimation in natural environments, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11214, 339–357.
- [49] **Wood, E., Baltrušaitis, T., Morency, L.P., Robinson, P. & Bulling, A.** (2016). Learning an appearance-based gaze estimator from one million synthesised images, *Eye Tracking Research and Applications Symposium (ETRA)*, 14, 131–138.
- [50] **Wojke, N., Hedrich, J., Droege, D. & Paulus, D.** (2016). Gaze-estimation for consumer-grade cameras using a Gaussian process latent variable model, *Pattern Recognition and Image Analysis*, 26(1), 248–255.
- [51] **Wu, Y.L., Yeh, C.T., Hung, W.C. & Tang, C.Y.** (2014). Gaze direction estimation using support vector machine with active appearance model, *Multimedia Tools and Applications*, 70(3), 2037–2062.
- [52] **Kacete, A., Séguier, R., Collobert, M. & Royan, J.** (2017). Unconstrained gaze estimation using random forest regression voting, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10113, 419–432.
- [53] **Tan, K.H., Kriegman, D.J. & Ahuja, N.** (2002). Appearance-based eye gaze estimation, *Proceedings of IEEE Workshop on Applications of Computer Vision*, 191–195.
- [54] **Noble, A.M., Miles, M., Perez, M.A., Guo, F. & Klauer, S.G.** (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems, *Accident Analysis & Prevention*, 151, 105959.

- [55] **Rakhmatulin, I. & Duchowski, A.T.** (2020). Deep Neural Networks for Low-Cost Eye Tracking, *Procedia Computer Science*, 176, 685–694.
- [56] **Turan, B., Eskikurt, H.I. & Can, M.S.** (2016). An application based on artificial neural network for determining viewpoint coordinates on a screen, *Elektronika Ir Elektrotechnika*, 22(2), 86–91.
- [57] **Thiyaneswaran, B. & Padma, S.** (2012). Iris Recognition using Left and Right Iris Feature of the Human Eye for Biometric Security System, *International Journal of Computer Applications*, 50(12), 37–41.
- [58] **Akram, A. & Debnath, R.** (2020). An automated eye disease recognition system from visual content of facial images using machine learning techniques, *Turkish Journal of Electrical Engineering and Computer Sciences*, 28(2), 917–932.
- [59] **Munson, M.C., Plewman, D.L., Baumer, K.M., Henning, R., Zahler, C.T., Kietzman, A.T., et al.** (2019). Autonomous early detection of eye disease in childhood photographs, *Science Advances*, 5(10).
- [60] **Viola, P. & Jones, M.** (2001). Rapid object detection using a boosted cascade of simple features, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1.
- [61] **Zhang, K., Zhang, Z., Li, Z. & Qiao, Y.** (2016). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks, *IEEE Signal Processing Letters*, 23(10), 1499–1503.
- [62] **Kazemi, V. & Sullivan, J.** (2014). One Millisecond Face Alignment with an Ensemble of Regression Trees, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp.1867–1874).
- [63] **Soetedjo, A.** (2012). Eye Detection Based-on Color and Shape Features, *International Journal of Advanced Computer Science and Applications*, 3(5).
- [64] **Majumder, G., Bhowmik, M.K. & Bhattacharjee, D.** (2013). Automatic Eye Detection Using Fast Corner Detector of North East Indian (NEI) Face Images, *Procedia Technology*, 10, 646–653.
- [65] **Knapik, M. & Cyganek, B.** (2021). Fast eyes detection in thermal images, *Multimedia Tools and Applications*, 80(3), 3601–3621.
- [66] **Gou, C., Wu, Y., Wang, K., Wang, K., Wang, F.Y. & Ji, Q.** (2017). A joint cascaded framework for simultaneous eye detection and eye state estimation, *Pattern Recognition*, 67, 23–31.
- [67] **Yu, M., Lin, Y. & Wang, X.** (2016). An efficient hybrid eye detection method, *Turkish Journal of Electrical Engineering and Computer Sciences*, 24(3), 1586–1603.
- [68] **Ghazali, K.H., Jadin, M.S., Jie, M. & Xiao, R.** (2015). Novel automatic eye detection and tracking algorithm, *Optics and Lasers in Engineering*, 67, 49–56.

- [69] **Doğan, F. & Türkoğlu, İ.** (2019). Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme, *DÜMF Mühendislik Dergisi*, 10(2), 409–445.
- [70] **Rosenblatt, F.** (1958). The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review*, 65(6), 386–408.
- [71] **Ivakhnenko, A. G. & Lapa, V.G.** (1966). *Cybernetic Predicting Devices*. CCM Information Corp., New York.
- [72] **Hinton, G.E. & Salakhutdinov, R.R.** (2006). Reducing the dimensionality of data with neural networks, *Science*, 313(5786), 504–507.
- [73] **Elmas, Ç.** (2011). *Yapay Zekâ Uygulamaları*. Türkiye: Seçkin yayıncılık.
- [74] **Şeker, A., Diri, B. & Balık, H.H.** (2017). Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme, *Gazi Journal of Engineering Sciences*, 3(3), 47–64.
- [75] **Schmidhuber, J.** (2015). Deep Learning in neural networks: An overview, *Neural Networks*, 61, 85–117.
- [76] **Zhang, X., Wang, L. & Su, Y.** (2020). Visual Place Recognition: A Survey from Deep Learning Perspective, *Pattern Recognition*, 107760.
- [77] **Zhang, Z., Zhao, Y., Liao, X., Shi, W., Li, K., Zou, Q., et al.** (2019). Deep learning in omics: A survey and guideline, *Briefings in Functional Genomics*, 18(1), 41–57.
- [78] **Cun, Y. le, Guyon, I., Jackel, L.D., Henderson, D., Boser, B., Howard, R.E., et al.** (1989). Handwritten Digit Recognition: Applications of Neural Network Chips and Automatic Learning, *IEEE Communications Magazine*, 27(11), 41–46.
- [79] **Hochreiter, S. & Schmidhuber, J.** (1997). Long Short-Term Memory, *Neural Computation*, 9(8), 1735–1780.
- [80] **Vincent, P., Larochelle, H., Bengio, Y. & Manzagol, P.A.** (2008). Extracting and composing robust features with denoising autoencoders, in: *Proceedings of the 25th International Conference on Machine Learning*, (pp.1096–1103).
- [81] **Salakhutdinov, R. & Hinton, G.** (2009). Deep Boltzmann machines, in: *Journal of Machine Learning Research*.
- [82] **Martens, J. & Sutskever, I.** (2011). Learning recurrent neural networks with Hessian-free optimization, in: *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*.
- [83] **Fukushima, K.** (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological Cybernetics*, 36(4), 193–202.

- [84] **LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., et al.** (1989). Handwritten Digit Recognition with a Back-Propagation Network, 396–404.
- [85] **LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P.** (1998). Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86(11), 2278–2324.
- [86] **Saraee, E., Jalal, M. & Betke, M.** (2020). Visual complexity analysis using deep intermediate-layer features, *Computer Vision and Image Understanding*, 195, 102949.
- [87] **Krizhevsky, A., Sutskever, I. & Hinton, G.E.** (2012). ImageNet classification with deep convolutional neural networks, *in: Adv Neural Inf Process Syst*, 1097–1105.
- [88] **Reyes, F., Fajardo, A. & Hernandez, A.** (2020). Convolutional neural network for automatic speech recognition of Filipino language, *International Journal of Advanced Trends in Computer Science and Engineering*, 9(1.1 Special Issue), 34–40.
- [89] **Dutta, P., Upadhyay, P., De, M. & Khalkar, R.G.** (2020). Medical Image Analysis using Deep Convolutional Neural Networks: CNN Architectures and Transfer Learning, *in: Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020, Institute of Electrical and Electronics Engineers Inc.*, (pp.175–180).
- [90] **Sultana, F., Sufian, A. & Dutta, P.** (2020). A review of object detection models based on convolutional neural network, *in: Advances in Intelligent Systems and Computing, Springer*, 1–16.
- [91] **Tsantekidis, A., Passalis, N., Tefas, A., Kannianen, J., Gabbouj, M. & Iosifidis, A.** (2020). Using Deep Learning for price prediction by exploiting stationary limit order book features, *Applied Soft Computing Journal*, 93.
- [92] **Liu, Z., Luo, P., Wang, X., Tang, X.** (2015). Deep learning face attributes in the wild, *In ICCV*, (pp.3730–3738).
- [93] **i·bug- resources - Facial point annotations.** Erişim: 8 Ağustos 2022, <https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>
- [94] **dlib C++ Library.** Erişim: 25 Temmuz 2022, <http://dlib.net/>
- [95] **Song, F., Tan, X., Liu, X. & Chen, S.** (2014). Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients, *Pattern Recognition*, 47(9), 2825–2838.
- [96] **Sahu, M. & Dash, R.** (2021). A survey on deep learning: Convolution neural network (cnn), *in: Smart Innovation, Systems and Technologies, Springer Science and Business Media Deutschland GmbH*, 317–325.
- [97] **Rachapudi, V. & Lavanya Devi, G.** (2020). Improved convolutional neural network based histopathological image classification, *Evolutionary Intelligence*, 1–7.

- [98] **Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R.R.** (2012). Improving neural networks by preventing co-adaptation of feature detectors.
- [99] **Hahn, S. and Choi, H.** (2020). Understanding dropout as an optimization trick, *Neurocomputing*, 398, 64–70.
- [100] **Home- OpenCV.** Erişim: 25 Temmuz 2022, <https://opencv.org/>
- [101] **Pedregosa Fabianpedregosa, F., Michel, V., Grisel Oliviergrisel, O., Blondel, M., Prettenhofer, P., Weiss, R., et al.** (2011). Scikit-learn: Machine learning in Python.
- [102] **Cai, H., Liu, B., Ju, Z., Thill, S., Belpaeme, T., Vanderborght, B., et al.** (2018). Accurate Eye Center Localization via Hierarchical Adaptive Convolution, *in: In Proceedings of the 29th British Machine Vision Conference, British Machine Vision Association, Newcastle*, 284.
- [103] **Kitazumi, K. & Nakazawa, A.** (2019). Robust Pupil Segmentation and Center Detection from Visible Light Images Using Convolutional Neural Network. *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, (pp.862–868).
- [104] **Choi, J.H., il Lee, K., Kim, Y.C. & Cheol Song, B.** (2019). Accurate Eye Pupil Localization Using Heterogeneous CNN Models, *Proceedings- International Conference on Image Processing, ICIP. 2019*, (pp.2179–2183), September.
- [105] **Larumbe-Bergera, A., Garde, G., Porta, S., Cabeza, R. & Villanueva, A.** (2021). Accurate pupil center detection in off-the-shelf eye tracking systems using convolutional neural networks, *Sensors*, 21(20).
- [106] **Ronneberger, O., Fischer, P. & Brox, T.** (2015). U-net: Convolutional networks for biomedical image segmentation, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351, 234–241.
- [107] **Villanueva, A., Ponz, V., Sesma-Sanchez, L., Ariz, M., Porta, S. & Cabeza, R.** (2013). Hybrid method based on topography for robust detection of iris center and eye corners, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 9(4).
- [108] **Jesorsky, O., Kirchberg, K.J. & Frischholz, R.W.** (2001). Robust face detection using the Hausdorff distance, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2091, 90–95.
- [109] **Zhang, W., Smith, M.L., Smith, L.N. & Farooq, A.** (2016). Eye center localization and gaze gesture recognition for human-computer interaction, *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 33(3), 314–325.

- [110] **Gou, C., Wu, Y., Wang, K., Wang, F.Y. & Ji, Q.** (2016). Learning-by-synthesis for accurate eye detection, *Proceedings- International Conference on Pattern Recognition*, 0, 3362–3367.
- [111] **Levinshtein, A., Phung, E. & Aarabi, P.** (2018). Hybrid eye center localization using cascaded regression and hand-crafted model fitting, *Image and Vision Computing*, 71, 17–24.
- [112] **Xiao, F., Huang, K., Qiu, Y. & Shen, H.** (2018). Accurate iris center localization method using facial landmark, snakuscule, circle fitting and binary connected component, *Multimedia Tools and Applications*, 77(19), 25333–25353.
- [113] **Xia, Y., Yu, H. & Wang, F.Y.** (2019). Accurate and robust eye center localization via fully convolutional networks, *IEEE/CAA Journal of Automatica Sinica*, 6(5), 1127–1138.
- [114] **Kim, S., Jeong, M. & Ko, B.C.** (2020). Energy Efficient Pupil Tracking Based on Rule Distillation of Cascade Regression Forest, *Sensors 2020*, 20(18), 5141.
- [115] **Lee, K. il, Jeon, J.H. & Song, B.C.** (2020). Deep Learning-Based Pupil Center Detection for Fast and Accurate Eye Tracking System, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12364 LNCS, 36–52.
- [116] **Fu, B. & Yang, R.** (2011). Display control based on eye gaze estimation, *Proceedings - 4th International Congress on Image and Signal Processing, CISP 2011, 1*, (pp.399–403).
- [117] **Utaminigrum, F., Ali Fauzi, M., Sari, Y.A., Primaswara, R. & Adinugroho, S.** (2016). Eye movement as navigator for disabled person, *ACM International Conference Proceeding Series*, (pp.1–5).
- [118] **Tao, L. & Changle, P.** (2008). Eye-gaze tracking research based on image processing, *Proceedings- 1st International Congress on Image and Signal Processing, CISP 2008, 4*, (pp.176–180).
- [119] **Park, S.-J. & Kim, B.-G.** (2020). Development of Low-Cost Vision-based Eye Tracking Algorithm for Information Augmented Interactive System, *Journal of Multimedia Information System*, 7(1), 11–16.
- [120] **Ari, A. & Hanbay, D.** (2019). Tumor detection in MR images of regional convolutional neural networks, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34(3), 1395–1408.
- [121] **Ari, A. & Hanbay, D.** (2018). An Expert Systems for Brain MR Images Classification by Using Ensemble Neural Network, *Anatolian Journal of Computer Sciences*, 3(2), 44–56.
- [122] **Gonzalez, R.C. & Woods, R.E.** (2007). Digital Image Processing, *Prentice-Hall, Inc. Upper Saddle River, NJ, USA*.

- [123] **Suzuki, S. & be, K.A.** (1985). Topological structural analysis of digitized binary images by border following, *Computer Vision, Graphics, and Image Processing*, 30(1), 32–46.
- [124] **Harish, A.B. & Prasad, A.R.** (2021). Automated 3D solid reconstruction from 2D CAD using OpenCV.
- [125] **Durna, Y. & Ari, F.** (2016). Polinom Fonksiyonları ile Göz Bakış Yeri Tespiti Geliştirilmesi ve Uygulaması, *Savunma Bilimleri Dergisi The Journal of Defense Sciences*, 15, 25–45.
- [126] **Morimoto, C.H. & Mimica, M.R.M.** (2004). Eye gaze tracking techniques for interactive applications, *Computer Vision and Image Understanding*, 98(1), 4–24.
- [127] **Lancaster, P. & Salkauskas, K.** (1986). Curve and surface fitting-an introduction.
- [128] **Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., et al.** (2016). Eye Tracking for Everyone, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 1063–6919).
- [129] **Sewell, W. & Komogortsev, O.** (2010). Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network, in: *Conference on Human Factors in Computing Systems- Proceedings*, (pp. 3739–3744).
- [130] **Xu, L.-Q., Machin, D. & Sheppard, P.** (1998). A Novel Approach to Real-time Non-intrusive Gaze Finding, in: *In Proc. BMVC*.
- [131] **Kim, M., Wang, O. & Ng, N.** (2017). Convolutional Neural Network Architectures for Gaze Estimation on Mobile Devices.
- [132] **Zhang, X., Sugano, Y., Fritz, M. & Bulling, A.** (2015). Appearance-based gaze estimation in the wild, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 4511–4520).
- [133] **Zhang, X., Sugano, Y., Fritz, M., & Bulling, A.** (2016). It's Written All Over Your Face: Full-Face Appearance-Based Gaze Estimation, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, (pp. 2299–2308).
- [134] **Abdelrahman, A.A., Hempel, T., Khalifa, A., & Al-Hamadi, A.** (2022). L2CS-Net: Fine-Grained Gaze Estimation in Unconstrained Environments.
- [135] **Jyoti, S. and Dhall, A.** (2018). Automatic Eye Gaze Estimation using Geometric Texture-based Networks, *Proceedings - International Conference on Pattern Recognition*, (pp. 2474–2479).
- [136] **Wong, E.T., Yean, S., Hu, Q., Lee, B.S., Liu, J. & Deepu, R.** (2019) Gaze Estimation Using Residual Neural Network, in: *2019 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops*.

- [137] **Chen, H.H., Hwang, B.J., Wu, J.S. & Liu, P.T.** (2020). The effect of different deep network architectures upon CNN-based gaze tracking, *Algorithms*, 13(5).
- [138] **Celebrating 20 years - Global leader in eye tracking - Tobii.** Eriřim: 15 Ağustos 2022, <https://www.tobii.com/>
- [139] **SR Research Ltd.- Eye-Tracking Company.** Eriřim: 8 Ağustos 2022, <https://www.sr-research.com/>
- [140] **Slanzi, G., Balazs, J.A. & Velásquez, J.D.** (2017). Combining eye tracking, pupil dilation and EEG analysis for predicting web users click intention, *Information Fusion*, 35, 51–57.
- [141] **Velásquez, J.D.** (2013). Combining eye-tracking technologies with web usage mining for identifying Website Keyobjects, *Engineering Applications of Artificial Intelligence*, 26(5–6), 1469–1478.
- [142] **Mora, K.A.F., Monay, F. & Odobez, J.M.** (2014). EYEDIAP: A database for the development and evaluation of gaze estimation algorithms from RGB and RGB-D cameras, *in: Eye Tracking Research and Applications Symposium (ETRA)*, (pp. 255–258).
- [143] **Sugano, Y., Matsushita, Y. & Sato, Y.** (2014). Learning-by-synthesis for appearance-based 3D gaze estimation, *in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 1821–1828).
- [144] **Kellnhofer, P., Recasens, A., Stent, S., Matusik, W., & Torralba, A.** (2019). Gaze360: Physically Unconstrained Gaze Estimation in the Wild, *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 6911–6920).
- [145] **Colaboratory'ye Hoř Geldiniz- Colaboratory.** Eriřim: 9 Ağustos 2022, <https://colab.research.google.com/>
- [146] **Janssen, R. G. J., Rothkrantz, L., Wiggers, I. P. & Geers, I.H.** (2010). Real Time Eye Blink Detection using a Configurable Processor, *Ph.D. Dissertation, Doctoral Dissertation, TU Delft, Delft University of Technology*.
- [147] **Ioffe, S. & Szegedy, C.** (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift, *in: 32nd International Conference on Machine Learning, ICML 2015*.
- [148] **Jie, H.J. & Wanda, P.** (2020). Runpool: A dynamic pooling layer for convolution neural network, *International Journal of Computational Intelligence Systems*, 13(1), 66–76.
- [149] **Akhtar, N. & Ragavendran, U.** (2020). Interpretation of intelligence in CNN-pooling processes: a methodological survey, *Neural Computing and Applications*, 32(3), 879–898.

- [150] **Sun, S., Hu, B., Yu, Z. & Song, X.** (2020). A stochastic max pooling strategy for convolutional neural network trained by noisy samples, *International Journal of Computers, Communications and Control*, 15(1).
- [151] **Nair, V. & Hinton, G.E.** (2010). Rectified Linear Units Improve Restricted Boltzmann Machines, *in: ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning*, (pp. 807–814).
- [152] **Kingma, D.P. & Ba, J.L.** (2015). Adam: A method for stochastic optimization, *in: 3rd International Conference on Learning Representations, ICLR 2015- Conference Track Proceedings*.
- [153] **Wilkinson, L. & Friendly, M.** (2009). History corner the history of the cluster heat map, *American Statistician*, 63(2), 179–184.
- [154] **de Silva, S., Dayarathna, S., Ariyaratne, G., Meedeniya, D., Jayarathna, S., Michalek, A.M.P., et al.** (2019). A Rule-Based System for ADHD Identification using Eye Movement Data, *MERCon 2019- Proceedings, 5th International Multidisciplinary Moratuwa Engineering Research Conference*, (pp.538–543).
- [155] **Elbaum, T., Braw, Y., Lev, A. & Rassovsky, Y.** (2020). Attention-Deficit/Hyperactivity Disorder (ADHD): Integrating the MOXO-dCPT with an Eye Tracker Enhances Diagnostic Precision, *Sensors (Basel, Switzerland)*, 20(21), 1–19.
- [156] **Cortes, C., Vapnik, V. & Saitta, L.** (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.

ÖZGEÇMİŞ

Ad-Soyad : **Kenan DONUK**

ÖĞRENİM DURUMU:

- **Lisans** : 2007, Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü
- **Yüksek Lisans** : 2014, Fırat Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü
- **Doktora** :

MESLEKİ DENEYİM:

- 2009-Devam Şırnak Üniversitesi – Öğretim Görevlisi

DOKTORA TEZİNDEN TÜRETİLEN ÇALIŞMALAR

SCI

1. **Donuk, K., Arı, A., Hanbay, D. (2022).** A CNN Based Real-Time Eye Tracker for Web Mining Applications. *Multimedia Tools and Applications*, 81:39103-39120, 2022.

TR DİZİN

1. **Donuk, K., Arı, A., Hanbay, D. (2021).** Yüz İmgelerinden Göz Bölgelerinin Tespitinde ESA Tabanlı Alternatif Bir Yaklaşım. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 2021. Cilt 32, Sayı 2, 735-743.

ULUSLARARASI KONFERANS (Bildiriler)

1. **Donuk, K., Hanbay, D. (2020, Ekim).** Video Based Real-Time Eye Tracking. 28th IEEE Signal Processing and Communications Applications (SIU-2020) (pp. 21-24).

ULUSLARARASI DERGİ

1. **Donuk, K., Hanbay, D. (2022).** Pupil Center Localization Based on Mini U-Net. *Journal of Computer Science*, 2022. IDAP-2022: International Artificial Intelligence and Data Processing Symposium, 185-191.