



İSKENDERUN TEKNİK

ÜNİVERSİTESİ

MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ

**YÜKSEK
LİSANS
TEZİ**

**GERÇEK ZAMANLI GÖRÜNTÜ
İŞLEME İLE SINIF ORTAMINDAKİ
ÖĞRENCİLERİN TANINMASI VE
DİKKAT DAĞINIKLIKLARININ TESPİTİ**

Mustafa Uğur UÇAR

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI**

HAZİRAN 2019



**GERÇEK ZAMANLI GÖRÜNTÜ İŞLEME ile SINIF ORTAMINDAKİ
ÖĞRENCİLERİN TANINMASI ve DİKKAT DAĞINIKLIKLARININ
TESPİTİ**

Mustafa Uğur UÇAR

**YÜKSEK LİSANS TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**İSKENDERUN TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ**

HAZİRAN 2019

Mustafa Uğur UÇAR tarafından hazırlanan “GERÇEK ZAMANLI GÖRÜNTÜ İŞLEME ile SINIF ORTAMINDAKİ ÖĞRENCİLERİN TANINMASI ve DİKKAT DAĞINIKLIKLARININ TESPİTİ” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile İskenderun Teknik Üniversitesi Elektrik Elektronik Mühendisliği Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Dr. Öğr. Üyesi Ersin ÖZDEMİR

Elektrik-Elektronik Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Başkan: Doç. Dr. Mustafa ORAL

Yazılım Anabilim Dalı, Çukurova Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Üye: Doç. Dr. Yakup KUTLU

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

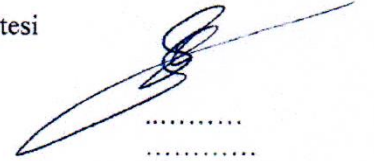
Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi Ersin ÖZDEMİR

Elektrik-Elektronik Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Tez Savunma Tarihi: 10/06/2019

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

Mühendislik ve Fen Bilimleri Enstitüsü Müdürü



ETİK BEYAN

İskenderun Teknik Üniversitesi Mühendislik ve Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez üzerinde Yükseköğretim Kurulu tarafından hiçbir değişiklik yapılamayacağı için tezin bilgisayar ekranında görüntülendiğinde asıl nüsha ile aynı olması sorumluluğunun tarafıma ait olduğunu,
- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.



Mustafa Uğur UÇAR

10/06/2019

GERÇEK ZAMANLI GÖRÜNTÜ İŞLEME ile SINIF ORTAMINDAKİ ÖĞRENCİLERİN
TANINMASI ve DİKKAT DAĞINIKLIKLARININ TESPİTİ (Yüksek Lisans Tezi)

Mustafa Uğur UÇAR

İSKENDERUN TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ

Haziran 2019

ÖZET

Tezin amacı, bir sınıf ortamında bulunan öğrencilerin kamera ile gerçek zamanlı olarak izlenmesi, kafalarının tespit ve takip edilmesi, yüz tanıma ile kim olduklarının tahmin edilmesi ve kafa yönlerinden hareketle dikkat dağınıklığı oranlarının tespitidir. Bu amaçla OpenCV ve Dlib görüntü işleme ve makine öğrenmesi kütüphaneleri kullanılarak C++ yazılım dili ile gerçek zamanlı çalışan bir uygulama geliştirilmiştir.

Uygulamanın yüz tanıma bölümünde dikkat dağınıklıkları incelenecek her bir öğrenci için yüz tanıma eğitim kümesi oluşturulabilmektedir. Kamera karşısındaki öğrenci değişik kafa pozisyonları ve mimikler yaparken uygulama her 5 saniyede bir öğrencinin yüzünün fotoğraflarını çekip kaydetmektedir. Yüz fotoğraflarını içeren bu eğitim kümesi kullanılarak Yerel İkili Örüntüler yöntemi ile yüz tanıma modeli oluşturulmaktadır.

Kafa yönünün ve dikkat dağınıklığının tespiti bölümünde ise her bir öğrenci için Perspective-n-Point problemi çözülerek öğrencilerin kafalarının x, y ve z eksenlerindeki dönme açıları (pitch, yaw ve roll) bulunmaktadır. Öğretmenle aynı doğrultuda bulunan bir kameraya doğru bakmakta oldukları varsayılan öğrencilerin kameraya doğru baktıklarında dikkatli oldukları kabul edilerek dikkat dağınıklıkları anlık olarak belirlenmektedir.

Yüz tanıma ve kafa yönü açılarının başarıları UPNA Head Pose Database ile test edilmiş, gerçekleştirilen testler sonrasında yüz tanımada en başarılı sonucu veren Yerel İkili Örüntüler yöntemi tercih edilmiştir. Görüntü işlemeye dayalı kafa yönü açılarını bulma yöntemi ise sensörlerle ölçülen açılara yakın sonuçlar vermiştir.

Dikkat dağınıklığı tespitinin başarısını ölçebilmek için yine UPNA Head Pose Database kullanılmıştır. Bu veritabanındaki videolardan çıkarılan rastgele 1000 fotoğraftan her biri 5 farklı kişi tarafından “Dikkatli” veya “Dikkatsiz” olarak işaretlenmiş, işaretlenen bu veri kümesinin yarısı eğitim yarısı da test kümesi olarak belirlenmiştir. Dikkat dağınıklığı sınıflandırıcıları etiketlenmiş veri kümesiyle eğitilmiştir. Sınıflandırıcı olarak Rastgele Orman, Karar Ağacı, DVM ve En Yakın Komşu makine öğrenmesi algoritmalarıyla denemeler yapılmıştır. En yüksek oran %72.4 doğrulukla DVM ile elde edilmiştir.

Anahtar Kelimeler : Gerçek zamanlı görüntü işleme, öğrencinin derse ilgisinin tespiti, görsel dikkat odağı tespiti, kafa yönü tespiti, yüz tanıma

Sayfa Adedi : 75

Danışman : Dr. Öğr. Üyesi Ersin ÖZDEMİR

RECOGNIZING STUDENTS' FACES IN THE CLASROOM AND ESTIMATING THEIR
ENGAGEMENT VIA COMPUTER VISION TECHNIQUES IN REAL TIME

(M. Sc. Thesis)

Mustafa Uğur UÇAR

ISKENDERUN TECHNICAL UNIVERSITY
ENGINEERING AND SCIENCE INSTITUTE

June 2019

ABSTRACT

Aim of this study is to track students in a classroom via webcam; detect, track and recognize their faces; and detect their engagement based on head pose estimation. For this purpose, we developed a real time C++ application by utilizing OpenCV and Dlib computer vision and machine learning libraries.

Face recognition module of our application allows us to create a face recognition training set for each student. Our application captures images of the student in front of the webcam in every 5 seconds while he/she is changing his/her head pose and facial expressions. Then, face recognition model is created via Local Binary Patterns Histograms method by using the face recognition training set.

In student engagement detection module, we solve Perspective-n-Point problem for each student to estimate their head rotation angles (pitch, yaw and roll). Assuming camera and the teacher are in the same direction, our application detects student engagement if students' head rotation points out the camera or not.

Success of face recognition and head pose estimation parts are tested with UPNA Head Pose Database. Local Binary Patterns Histograms method is preferred in our application, which gave the best accuracies. Head pose estimation tests showed that the angles returned from our computer vision based method are nearly the same as the angles measured by sensors.

Student engagement detection part is also tested with UPNA Head Pose Database. We extracted random 1000 images from the database. 5 human labelers annotated each image as "Engaged" or "NotEngaged". Half of the annotated images are used for training and others are used for testing. Engagement classifiers were trained with the labelled dataset. Experiments were made with Random Forest, Decision Tree, SVM and KNN machine learning algorithms to classify student engagement. SVM classifier achieved 72.4% accuracy on the labelled dataset.

Key Words : Real-time image processing, student engagement, visual focus of attention,
head pose estimation, face recognition
Page Number : 75
Supervisor : Assist. Prof. Dr. Ersin ÖZDEMİR

TEŞEKKÜR

Tez konumun şekillenmesinde yardımını esirgemeyen, beni motive eden danışman hocam Sn. Dr. Öğr. Üyesi Ersin ÖZDEMİR'e; eğitim ve öğretim hayatım boyunca bilgilerinden istifade ettiğim tüm hocalarıma; yorucu ve yıpratıcı geçen tez hazırlama sürecinde bana her türlü desteği sağlayan, hep yanımda olan, moral kaynağım, canım eşim Berna'ya; benimle üzüлüp benimle sevinen, üzerimde tarifsiz emekleri olan, haklarını ödeyemeyeceğim başta sevgili annem ve babam olmak üzere tüm aileme çok teşekkür ederim.

Bu tez çalışmamı, kendime örnek aldığım kişilerin başında gelen, hayatımda önemli yeri olan kıymetli dedem İvriz Köy Enstitüsü mezunu emekli öğretmen Mehmet Emin Yıldız'a ithaf ediyorum.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ	ix
ŞEKİLLERİN LİSTESİ.....	x
RESİMLERİN LİSTESİ.....	xi
SİMGELER VE KISALTMALAR	xii
1. GİRİŞ	1
2. ÖNCEKİ ÇALIŞMALAR	5
2.1. Dikkat Dağımlılığının ve Bakış Yönünün Tespiti ile İlgili Çalışmalar.....	5
2.2. Yüz Tanıma ile İlgili Çalışmalar.....	14
3. MATERYAL ve YÖNTEM	24
3.1. Donanım.....	24
3.1.1 Web kamerası.....	24
3.2. Yazılım Geliştirme Ortamı ve Yazılım Kütüphaneleri	25
3.2.1. Qt Creator.....	25
3.2.2. OpenCV	26
3.2.3. Dlib	27
3.3. Kullanılan Yöntemler.....	28
3.3.1. Yönlendirilmiş Gradyanların Histogramı (YGH).....	29
3.3.2. Yerel İkili Örüntüler (YİÖ).....	35
3.3.3. Destek Vektör Makineleri (DVM).....	38
3.3.4. PnP probleminin çözülmesi ve Euler açılarının bulunması	42

	Sayfa
3.3.5. Göz kısıklığının tespit edilmesi	44
3.4. Kullanılan Veri Kümesi	45
3.4.1. UPNA Head Pose Database	45
4. DİKKAT DAĞINIKLIĞI TESPİT SİSTEMİ (DDTS) YAZILIMI	47
4.1. Yüz tanıma Modeli Oluşturma Bölümü	48
4.2. Programın Ana Bölümü	52
5. ARAŞTIRMA BULGULARI ve TARTIŞMA	59
5.1. Kafa Yönü Tahmini ile İlgili Bulgular	59
5.2. Yüz Tanıma ile İlgili Bulgular	60
5.3. Sistemin Geneli ile İlgili Bulgular	61
6. SONUÇ ve ÖNERİLER	64
KAYNAKLAR	66
ÖZGEÇMİŞ	74
DİZİN	75

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. 2001-2005 yılları arasında yapılmış olan Harris-Westin anket sonuçları	18
Çizelge 3.1. Kullanılan donanım.	24
Çizelge 3.2. Yazılım geliştirme ortamı ve kullanılan yazılım kütüphaneleri.	25
Çizelge 5.1. Uygulamanın kafa yönü tahmini kısmının başarısı.	59
Çizelge 5.2. Uygulamanın yüz tanıma kısmının başarısı.....	60
Çizelge 5.3. DDTS'nin genel başarısı – 1. senaryo.	61
Çizelge 5.4. DDTS'nin genel başarısı – 2. senaryo (hesaplanan Euler açıları ile).	63
Çizelge 5.5. DDTS'nin genel başarısı – 2. senaryo (ölçülen Euler açıları ile).	63

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. İtici etki, çekici etki	7
Şekil 2.2. Geleneksel yöntemlerle biyometriye dayalı kimlik doğrulaması	16
Şekil 2.3. Biyometrik karakteristik örnekleri.....	17
Şekil 2.4. ILSVRC yarışmasındaki yıllara göre Top-5 hata oranları.....	23
Şekil 3.1. Gradyanın yönü ve gradyanın büyüklüğü.....	30
Şekil 3.2. x ve y eksenindeki gradyanların elde edilmesinde kullanılan maskeler.....	30
Şekil 3.3. Her bir 8 x 8 piksellik görüntü bloğu için gradyanların çıkarılması.....	32
Şekil 3.4. Gradyanların histogramının oluşturulması	33
Şekil 3.5. Gradyanların histogramı	34
Şekil 3.6. Yerel ikili Örüntüler yönteminin uygulanması.....	36
Şekil 3.7. Dairesel YİÖ örnekleri.	36
Şekil 3.8. YİÖ histogramının oluşturulması	37
Şekil 3.9. DVM’de sınıfları ayırmak için çizilen çizgiler.....	39
Şekil 3.10. DVM’de karar sınırı ve destek vektörleri	39
Şekil 3.11. DVM’de karar sınırının matematiksel ifadesi	40
Şekil 3.12. DVM’de çekirdek hilesinin uygulanması.....	41
Şekil 3.13. Euler açıları.....	43
Şekil 3.14. Gözün açık ve kapalı olma durumuna göre nirengi noktalarının durumu ...	44
Şekil 4.1. DDTS’nin yüz tanıma modeli oluşturma bölümünün akış diyagramı.....	49
Şekil 4.2. DDTS yazılımı yüz tanıma modeli oluşturma	54
Şekil 5.1. DDTS’nin başarısını tespit için kullanılan veri kümesinden örnek görüntüler .	62

RESİMLERİN LİSTESİ

Resim	Sayfa
Resim 2.1. Wollaston etkisi	7
Resim 3.1. Logitech HD Webcam C525	25
Resim 3.2. Orijinal görüntü ile x, y eksenindeki gradyan görüntüler.....	31
Resim 3.3. YİÖ yönteminin ışıklandırma değişimi karşısında başarısı.....	35
Resim 3.4. UPNA Head Pose Database'den alınan örnek görüntüler	46
Resim 4.1. DDTS yazılımının ekran görüntüsü.....	47
Resim 4.2. DDTS yazılımı yüz tanıma modeli oluşturma ekran görüntüsü	50
Resim 4.3. DDTS yazılımı yüz tanıma modeli oluşturma – Örnek fotoğraflar.....	51
Resim 4.4. DDTS yazılımı yüz tanıma modeli oluşturma	51
Resim 4.5. DDTS yazılımı ana bölümü - ayarlamalar.....	53
Resim 4.6. DDTS yazılımı ana bölümü – çeşitli ekran görüntüleri.....	55
Resim 4.6. (Devam) DDTS yazılımı ana bölümü – çeşitli ekran görüntüleri.....	56
Resim 4.7. DDTS yazılımı ana bölümü – anlık toplam dikkat dağınıklığı sonuçları	57
Resim 4.8. DDTS yazılımı ana bölümü – dikkat dağınıklığı tespit sonuçları	58

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

mm

milimetre

ms

milisaniye

°

derece

Kısaltmalar

Açıklamalar

DDTS

Dikkat Dağınıklığı Tespit Sistemi

YİÖ

Yerel İkili Örüntüler

DVM

Destek Vektör Makineleri

TBA

Temel Bileşenler Analizi

DAA

Doğrusal Ayırtaç Analizi

ESA

Evrişimsel Sinir Ağları

ILSVRC

ImageNet Large Scale Visual Recognition Challenge

KNN

k-Nearest Neighbors

1. GİRİŞ

Teknolojik gelişmelerle birlikte bilgisayarlardaki işlemcilerin güçlenmesi ve işlem kapasitelerinin artmasıyla, büyük verilerin/görüntülerin gerçek zamanlı olarak işlenmesi mümkün hâle gelmiş ve bu sayede görüntü işleme alanındaki çalışmalar hızla artmaya başlamıştır.

İnsan-bilgisayar etkileşiminin giderek önem kazanması ve bilgisayarların/akıllı aygıtların hayatımızın her alanına girmesi, bilgisayarlara insansı yeteneklerin kazandırılması ihtiyacını da beraberinde getirmiştir. Bu amaçla yapılan çalışmalar özellikle görüntü işleme, bilgisayar görmesi (computer vision) alanında yoğunlaşmıştır.

Kameradan alınan görüntülerin bilgisayarlarda işlenmesi ile:

- objelerin tanınması (object recognition),
- yüz bulma (face detection),
- yüz takibi (face tracking),
- yüz tanıma (face recognition),
- duygu tanıma (emotion recognition),
- hareket takibi ve tanıma (human action recognition and tracking),
- cinsiyet/yaş tespiti (gender/age classification),
- işaret dili algılama (gesture recognition for sign language),
- sürücü uykululuğunun tespiti (driver drowsiness detection),
- bakış yönü tespiti (gaze detection),
- kafa yönü tahmini (head pose estimation)

gibi konular araştırmacıların ilgisini çeken ve sıkça çalışılmakta olan konulardandır.

İnsan davranışlarının görüntü işleme teknikleri ile tespit edilip analiz edilmesi konusu da son zamanlarda popülerlik kazanmış ve farklı alanlarda önemli çalışmalara zemin hazırlamıştır:

- Otomatik aldatma tespiti (Wu, Singh, Davis ve Subrahmanian, 2018),
- Yüz ifadesi tanıma (Moore ve Bowden, 2011),
- Şüpheli insan davranışlarının tespiti (Tripathi, Jalal ve Agrawal, 2017),
- Yayaların analizi (Keller ve Gavrilu, 2014),
- İşaret dili için ifade tanıma (Arı, Uyar ve Akarun, 2008),

- Sürücü uykululuğunun tespiti (Zhang, Cheng ve Lin, 2012),
- Sporcu performans analizi (Bertasius, Park, Yu ve Shi, 2017),
- Oyuncu hareketlerinin önceden tahmin edilmesi (Felsen, Agrawal ve Malik, 2017),
- İnsan vücut hareketlerinin önceden tespit edilmesi ve tanınması (Wang ve Hoai, 2018)

bu alanlardan ve çalışmalardan sadece birkaçıdır.

Eğitim alanında görüntü işleme ile yapılan çalışmaların ise çok yaygın olmamakla birlikte (Sümer ve diğerleri, 2018) son yıllarda artış gösterdiği gözlemlenmektedir.

Bizim bu çalışmadaki yaklaşımımız, insan etkileşimlerinin güncel teknolojileri kullanarak (konuşma tanıma, görüntü işleme, makine öğrenmesi gibi) otomatik biçimde analiz edilmesi olarak tarif edilen sosyal sinyal işleme (social signal processing) dayanmaktadır.

Sınıf eğitimlerindeki en önemli iki aktör öğretmen ve öğrencidir. Eğitimin başarısı temelde bu iki aktöre bağlıdır. Bu çalışmanın motivasyon kaynağı, eğitimdeki başarıyı artırabilmek adına öğretmenin öğrenciye ulaşamadığı anları tespit edebilecek bir indikatörün tespit edilip edilemeyeceği olmuştur. Öğrencilerin ders sırasındaki sözsüz davranışlarının görüntü işleme teknikleri ile saptanması ve analiz edilmesi hedeflenmiştir.

Ders esnasında öğretmenin öğrenciye ulaşabildiğini gösteren en önemli indikatör öğrencinin dikkatini öğretmene ve işlenen konuya vermesidir diyebiliriz. Öğrencinin dikkatini öğretmene ve derse vermesi öğrenci başarısında, dersin verimli olmasında ve öğretmen motivasyonunda önemli bir role sahiptir. Öğrencinin görsel dikkatini derse ve öğretmene verdiği en önemli göstergelerinden birisi de öğrencinin öğretmene doğru bakıyor olmasıdır denilebilir.

Bu çalışmada, web kamerası, görüntü işleme teknikleri ve makine öğrenmesi algoritmaları kullanılarak sınıf ortamında bulunan öğrencilerin gerçek zamanlı olarak tanınması, gerçek zamanlı olarak dikkat dağınıklığı oranlarının bulunması ve öğretmene anlık geribildirimler sunularak ders akışının iyileştirilmesi hedeflenmiş ve bu amaçla da bir uygulama geliştirilmiştir. Ayrıca öğrencilerin dikkat dağınıklığı oranlarının tahmin edilmesi vasıtasıyla:

- Öğrenim tecrübesinin ve farkındalığının arttırılması,
- Öğrencilerin dikkatinin düştüğü anlarda öğretmene müdahale imkânı vermesi,
- Geliştirilen uygulamadan alınan geribildirimler doğrultusunda öğrenim materyallerinin ilgiyi artıracak şekilde güncellenmesi

hedeflenmiştir.

Çalışmayla ilgili detaylara geçmeden önce dikkatin, dikkat dağınıklığının ne olduğuna ve dikkat dağınıklığının belirtilerine değinmek faydalı olacaktır.

Türk Dil Kurumu Güncel Türkçe Sözlük'te dikkat, "duygularla düşünceyi bir şey üzerinde toplama, uyanıklık" olarak tanımlanmaktadır. Dikkat dağınıklığı ise dikkati belli bir işe/kişiye verememe, yoğunlaşamama veya dikkati sürdürmemeye durumudur. Günümüzde çocuklarda ve erişkinlerde yaygın olarak farklı seviyelerde görülmektedir.

Bireylerdeki dikkat eksikliğinin gözlemlenebilir belirtileri arasında:

- konuşana ya da odaklanması gereken nesneye değil başka yöne bakmak,
- başka işlerle uğraşmak,
- gözlerini kapamak, uyuklamak,
- söylenenleri algılamamak

sayılabilir.

Bu belirtiler insanlar tarafından kolaylıkla algılanabilmekte iken bunların bilgisayarlar tarafından tespit edilmesi, üzerinde hâlen çalışılmakta olan, güncelliğini koruyan bir konudur. Yapılan çalışmalarda ise dikkat dağınıklığı genel olarak:

- görüntü işleme teknikleri ya da ticarî cihazlar kullanılarak kafa yönünün, göz hareketlerinin takip edilmesi suretiyle bakış yönünün belirlenmesi ve bakışın dikkatin verilmesi gereken noktaya dönük olup olmamasına göre,
- kişinin kafasına takılan bir cihaz yardımıyla EEG sinyallerinin ölçülmesi ve işlenmesi yoluyla,

tespit edilmeye çalışılmıştır:

Kolay uygulanabilirliğinden ve maliyetinin düşük olmasından ötürü bu çalışmada görüntü işleme tekniklerinin kullanılması tercih edilmiştir. Öğrencilerin web kamerası ile tek tek

değil grup olarak izlenmesinden ötürü bakış yönlerinin tespitinde öğrencilerin gözlerinin yönünü takip etmek mümkün olmamıştır. Kafa yönünün bakış yönü ile genelde aynı olmasından ve kafa yönünün dikkat dağınıklığının tespitinde başarılı sonuçlar vermesinden ötürü bu çalışmada kafa yönüne dayalı olarak dikkat dağınıklığını tespit edecek bir yazılımın geliştirilmesine karar verilmiştir.

Kullanılan ölçüt kafa hareketlerinin görüntü işleme teknikleriyle tahmin edilmesi ile elde edilmektedir. Bu sayede fiziksel teması gerektiren herhangi bir cihazın kullanılmasına da gerek olmamıştır.

Tezin amaçlarını:

1. Öğrencilerin sınıf ortamındaki görüntüsünü web kamerasından alıp gerçek zamanlı olarak işleyecek olan bir uygulama geliştirilmesi,
2. Geliştirilecek uygulama vasıtasıyla öğrencilerin fotoğraflarının çekilerek yüz tanıma modellerinin oluşturulması ve bu modeller kullanılarak öğrencilerin tanınması,
3. Her bir öğrencinin yüzlerinin ayrı ayrı tespit ve takip edilmesi,
4. Yüzlere ait özniteliklerin çıkarılması ve bunların kafa yönü bulma aşamasında kullanılması,
5. Kafa yönlerinin durumuna göre dikkat dağınıklığı oranlarının her bir öğrenci için ve sınıf ortalaması olarak belirlenmesi

olarak sıralayabiliriz.

2. ÖNCEKİ ÇALIŞMALAR

2.1. Dikkat Dağınıklığının ve Bakış Yönünün Tespiti ile İlgili Çalışmalar

Başka bir kişinin nereye baktığının doğru bir şekilde tespit edilmesi günlük etkileşim ve iletişimde can alıcı bir fonksiyondur (Argyle ve Cook, 1976; Emery, 2000). Genelde, gözlemcilerin başka kişilerin bakış yönlerinin tespitinde gayet başarılı olduğu yapılmış olan çeşitli çalışmalarda belirtilmiştir (Gibson ve Pick, 1963; Symons, Lee, Cedrone ve Nishimura, 2004; Bock, Dicke ve Thier, 2008).

Görsel dikkatin tespit edilmesinde kullanılan en önemli ölçütlerin başında da bakış yönü (gaze direction) gelmektedir. Bakış yönünün dikkat seviyesinin tespit edilmesindeki rolü 1970'lerde Argyle ve Cook (1976) tarafından yapılan çalışmada vurgulanmış olup, sonradan yapılan çalışmalarda da bu tespitler doğrulanmıştır (Vertegaal, Slagter, Veer ve Nijholt, 2001).

Gözlerin uzayda işaret ettiği doğrultu olarak tanımlanan bakış yönünün kafa yönü (head orientation) ve göz yönünün (eye orientation) toplamından oluştuğu belirtilmiştir. Yani, gerçek bakış yönünün tespit edilebilmesi için her iki yönün bilinmesi gerekmektedir (Stiefelhagen ve Zhu, 2002).

Stiefelhagen ve Zhu (2002) tarafından yapılan çalışmada:

1. Toplantıdaki katılımcıların bakış yönlerine kafa yönlerinin etkisi ne kadardır?
2. Kafa yönü temel alınmak suretiyle kişilerin görsel dikkat odağını ne derece doğru tahmin edebiliriz?

sorularına yanıt verebilmek için deneyler yapılmıştır. Bu kapsamda 4 kişiden oluşan bir yuvarlak masa toplantısı organize edilmiş, her bir katılımcı için 10'ar dakikalık oturumlar düzenlenerek oturumlardaki veriler toplanmıştır. Her bir katılımcı için ayrı ayrı yapılan oturumlarda ilgili katılımcının ISCAN takip cihazını ve kamerayı kafasına takması istenmiş ve 10 dakika boyunca katılımcının toplantı sırasındaki kafa/göz/bakış yönü, göz kırpma bilgileri 1 dereceden az bir sapmayla tespit ve takip edilmiştir. Toplanan verilerin analiz edilmesiyle şu sonuçlara ulaşılmıştır:

1. Çoğu zaman kişiler dikkatlerini odakladıkları tarafa bakmak için gözlerini ve kafalarını aynı yöne çevirmişlerdir. Öznelerin göz ve kafalarının aynı yöne bakma oranı %87 olarak bulunmuştur.
2. Kişilerin bakış yönlerini değiştirmek için kafalarının yönünü değiştirmesi kişiden kişiye çeşitlilik göstermektedir. 2. kişi için bu oran %53 iken 4. kişi için %96 çıkmış; ortalama oran ise %68.9 olarak bulunmuştur.
3. Kafa yönünün bakış yönüne katkı oranı en düşük olan 2. kişide bile bu oran %50'den fazla çıkararak bakış yönüne yarıdan fazla katkıyı sağlamıştır.
4. Toplam görüntülerin %21.7'si göz kırpma anlarına denk gelmektedir. Bu da ISCAN gibi hata oranı düşük ve ticari ürünlerde bile görüntülerin beşte birinde göz yönünün ve dolayısıyla bakış yönünün düzgün tespit edilemediğini göstermektedir.

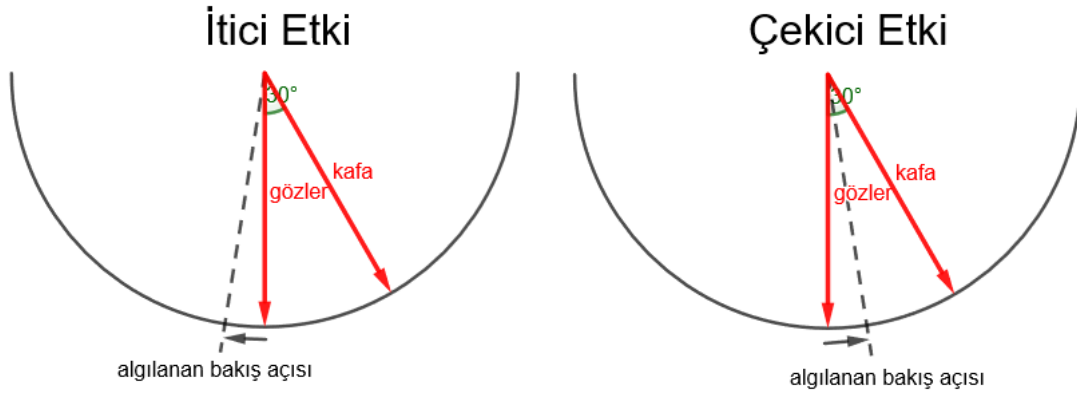
Özetlemek gerekirse, kafa yönünün bakış yönüne ortalama %68.9, dikkat odağına ise %88.7 oranında katkı sağladığı ortaya konulmuştur. Bu da kafa yönünün dikkat dağınıklığının tespitinde iyi bir gösterge olduğunu ortaya çıkarmıştır (Stiefelbogen ve Zhu, 2002).

Başka bir deyişle, bir kişinin görsel dikkat odağı (VFOA-visual focus of attention) onun bakış yönü ve dolayısıyla kafa yönü ile yakından ilgilidir diyebiliriz. Zira Ba ve Odobez (2008) de yapmış oldukları çalışmada, toplantılardaki katılımcıların görsel dikkat odaklarının tespitinde kafa yönlerinden faydalanmıştır.

Kafa yönünün algılanan bakış yönüne önemli etkisi olduğu birçok çalışmada vurgulanmakla birlikte çoğu çalışmada bu etkinin mahiyetinden bahsedilmemektedir. Bu etkinin mahiyetine dair yapılan çalışmalarda ise kafa ve göz yönleri aynı doğrultuda olmadığı zaman algılanan bakış yönünde iki çeşit sapma olduğu belirtilmiştir (Langton, Watt ve Bruce, 2000; Klutzz, Mayes, West ve Kerby, 2009; Otsuka, Mareschal, Calder ve Clifford, 2014; Moors, Verfaillie, Daems, Pomianowska ve Germeys, 2016). Literatürde geçen bu iki sapma:

1. İtici etki (repulsive effect, overshoot effect): Algılanan bakış yönünün kafa yönünün tersine bir doğrultuda olması
2. Çekici etki (attractive effect, towing effect): Algılanan bakış yönünün kafa yönü ile göz yönü arasında olması

olarak tarif edilmiştir. Çekici etki ve itici etkinin algılanan bakış yönüne etkisi Şekil 2.1'de görülmektedir.



Şekil 2.1. İtici etki (solda), çekici etki (sağda)

Wollaston (1824) tarafından yapılmış olan çalışma bu alandaki öncü çalışmalardandır. Wollaston gözleri karşıya, kafası ise bize göre sola doğru yönelmiş olan bir portreden hareketle, gözleri bir önceki resimden kopyalanan, kafası ise bize göre sağa yönelmiş bir portre çizmiş ve bu iki portredeki bakış yönlerini incelemiştir. İki portredeki gözler de birebir aynı olmasına rağmen ikinci portredeki kişi sağa doğru bakıyor olarak algılanmaktadır. Bu çalışmadan ötürü, bakış yönünün kafa yönünün olduğu tarafa doğru olarak algılanması çekici etkinin yanı sıra Wollaston etkisi (Wollaston effect) olarak da adlandırılmaktadır (Otsuka ve diğerleri, 2014). Bahsedilen etki Resim 2.1’de görülmektedir.



Resim 2.1. Wollaston etkisi (Otsuka ve diğerleri, 2014)

1960'larda bu konuyla ilgili üç önemli çığır açan makale yayınlanmıştır (Gibson ve Pick, 1963; Cline, 1967; Anstis, Mayhew ve Morley, 1969). Bu çalışmalardan ikisinin (Gibson ve Pick, 1963; Cline, 1967) sonuçları birçok yazar tarafından incelenmiş ve farklı şekillerde yorumlanmıştır. Moors ve diğerleri (2016) tarafından yapılan çalışmada ise bu iki çalışmanın

farklı şekillerde yorumlanmasına neden olan karışıklığın muhtemel sebepleri tartışılmıştır. Cline (1967)'nin sonuçları diğer yazarlar tarafından tartışma konusu olmuş olmasına rağmen net bir şekilde çekici etkiyi işaret etmekte iken Gibson ve Pick (1963)'in sonuçlarının literatürde sorgulanmamış olmasına rağmen en belirsiz olarak görüldüğü ifade edilmiştir (Moors ve diğerleri, 2016). Moors ve diğerleri (2016) bu belirsizlikleri gidermek ve konuyu açıklığa kavuşturmak adına her iki çalışmada yapılmış olan deneyleri tekrarlamış ve her iki çalışmanın sonucunda da itici etkinin mevcut olduğunu tespit etmişlerdir.

Farklı yorumlar ve çalışmalar olsa da kafa yönünün bakış yönüne önemli oranda etkisi olduğu rahatlıkla söylenebilir.

Kafa yönü tespiti (head pose estimation) kafaya takılan ticarî yazılımlarla yapılabildiği gibi kafaya herhangi bir cihaz takmaksızın kameradan alınan görüntülerin işlenmesi yöntemiyle de yapılabilmektedir. Bu yöntem yüksek çözünürlükte görüntülere ihtiyaç duymamakta ve maliyet yönünden de tercih edilmektedir.

Düşük çözünürlüklü videolarda göz konumunun bulunması mümkün olmadığından, Yücel, Salah, Ç. Meriçli ve T. Meriçli (2010) tarafından yapılan çalışmada görüntü işleme teknikleri kullanılmış, eliptik silindir tabanlı kafa duruşu çözümlemesiyle elde edilmiş kestirimler bakış yönüne dönüştürülmüş ve bu dönüşüm için bir Gauss süreci bağlantımı kullanılması önerilmiştir.

Gee ve Cipolla (1994) da yaptıkları çalışmada görüntü işleme tekniklerini tercih etmiştir. Kullandıkları yöntemde önce yüzdeki beş nokta (sağ göz, sol göz, ağzın sağ ve sol bitim noktaları ve burun ucu) tespit edilmiştir. Göz ve ağız noktalarından hareketle de burnun ağız üstündeki başlangıç kısmının bulunarak bu noktadan burun ucuna çizilen bir doğru (facial normal) ile bakış yönü kabul edilebilir doğrulukta bulunmuştur. Çok işlem gücü gerektirmeyen bu yöntem video görüntülerinde de gerçek zamanlı olarak uygulanmış ve elde edilen sonuçlar paylaşılmıştır.

Sapienza ve Camilleri (2011) ise Gee ve Cipolla (1994)'nin yöntemini biraz daha geliştirmiş, ağız için kullanılan iki nokta yerine ağzın orta noktasını almak suretiyle 4 noktadan yararlanmıştır. Viola-Jones algoritmasını kullanarak yüz ve yüzdeki öznitelikler tespit edilmiş, normalize edilmiş kareler farkının toplamı (normalized sum of squared difference)

yöntemi ile yüz öznitelikleri takip edilmiş, geometrik bir yaklaşım kullanılarak takip edilen özniteliklerden kafa yönü tahmin edilmiştir. Tasarladıkları bu kafa yönü tahmini sistemi kafa yönünü yaklaşık 6 derece sapma ile gerçek zamanlı olarak (2 ms) bulabilmiştir.

Zaletelj ve Kosir (2017) tarafından yapılan çalışmada Kinect kameralar ile sınıf içerisindeki öğrenciler izlenmiş, onların vücut duruşları, bakış yönleri, yüz ifadeleri belirlenmiş, görüntü işleme vasıtasıyla öznitelikler çıkarılmış, makine öğrenme algoritmaları kullanılarak öznitelik setlerinden oluşan 3 ayrı veri kümesi ile 7 ayrı sınıflandırıcı eğitilmiştir. 5 adet gözlemci tarafından sınıftaki öğrencilerin görüntülerindeki her bir saniye için dikkat seviyelerinin tahmin edilmesi istenmiş ve bu veri de veri kümelerine dâhil edilerek eğitim aşamasında referans olarak kullanılmıştır. Eğitilen model 18 kişilik bir veri kümesi ile test edilmiş ve dikkat dağınıklığının otomatik tespitinde %75.3'lük bir başarı göstermiştir.

Konuyla ilgili olarak "Sınıf içerisindeki öğrencilerin dikkatlerinin kamera tabanlı tahmini" başlıklı doktora tezinde ise:

- Sınıf içerisindeki öğrenciler birden çok kamera ile izlenerek öğrencilerin kafa yönlerinin yatay düzlemdeki hareketleri (yaw) tespit edilmiş,
- Sınıfın arka kısmında bulunan ve öğretmenin sınıf tahtası boyunca tek bir düzlemde olan hareketleri izlenerek yatay düzlemdeki pozisyonu belirlenmiş,
- Her bir öğrenciye ders esnasındaki her 10 dakikalık periyottaki dikkat seviyelerini raporlamaları amacıyla anket düzenlenmiş
- Anket cevaplarına göre dikkatli öğrencilerin yatay ekseninde kafa yönleri öğretmenin bulunduğu yöne doğru olan öğrenciler olduğu tespit edilmiştir (Raca, 2015).

Krithika ve Lakshmi Priya (2016) e-öğrenme kapsamında eğitim içeriklerinin iyileştirilebilmesi amacıyla öğrencinin kamerayla izlendiği bir sistem geliştirmiştir. Bu sistem kameradan alınan görüntüyü işlemekte, yüz ve gözlerin tespiti için Viola-Jones ve Yerel İkili Örüntüler (Local Binary Patterns) algoritmalarını kullanmaktadır. Öğrencinin kafasının ve gözlerinin tespit edilip edilememesine göre de konsantrasyon seviyesini belirlemeye çalışmaktadır. Yüksek, orta ve düşük olmak üzere üç ayrı konsantrasyon seviyesi belirlenmiş olup yüzün tespit edilemediği görüntü kareleri düşük konsantrasyon, kafanın tam olarak karşıya bakmadığı durumlarda yüzün tespit edilip gözlerden birinin tespit edilemediği kareler orta konsantrasyon, yüzün ve gözlerin tespit edildiği kareler ise yüksek

konsantrasyon olarak işaretlenmiştir. Geliştirilen sistem 5 öğrenciye 10 dakikalık bir video kursu izletilmek suretiyle test edilmiş ve test sonuçları paylaşılmıştır. Test sonuçlarına göre kurs esnasında dikkatin azaldığı ve iyileştirilmesi gereken bölümler tespit edilmiştir.

Ayvaz ve Gürüler (2017) de sınıf içi eğitimin kalitesini artırmayı amaçlayan bir çalışma yapmıştır. Dikkat seviyesi tespitine benzer olarak bu çalışmada da öğrencilerin kamera ile izlenerek duygusal durumlarının gerçek zamanlı olarak tespit edilmesi hedeflenmiştir. OpenCV ve dlib görüntü işleme kütüphanelerini kullanarak geliştirdikleri bu sistem, Viola-Jones algoritması ile yüz tespiti yapmakta, sonrasında yüzdeki özniteliklere göre duygu durumu tespiti yapmaktadır. Duygu durumu tespitinde kullanılan eğitim veriseti 12 lisans öğrencisinin (5 kız, 7 erkek) yüz verilerinden elde edilmiştir. Her bir öğrencinin 7 farklı duygusal durum(mutluluk, üzüntü, öfke, şaşkınlık, korku, tikslenme ve normal) için yüz öznitelikleri bu eğitim verisetine kaydedilmiştir. Bu veriseti ile eğitilen sistemde destek vektör makineleri ile öğrencilerin ders boyunca ağırlıklı duygusal durumları %97.15 doğrulukla tespit edilmiştir. Bu sistem sayesinde eğitimci tarafından, gerekli durumlarda öğrencilerle doğrudan iletişime geçilmesi ve motivasyonu arttırıcı yönlendirmeler yapılması hedeflenmiştir.

Yukarıdaki çalışmaların yanı sıra, bakış yönüne dayalı olarak öğrencinin dikkatinin, derse katılımının görüntü işleme teknikleri ile otomatik olarak belirlenmesi konusunda yapılan birçok çalışma (Bidwell ve Fuchs, 2011; D'Mello, Olney, Williams ve Hays, 2012; Asteriadis, Tzouveli, Karpouzis ve Kollias, 2009) vardır. Whitehill, Serpell, Lin, Foster ve Movellan (2014) öğrenci dikkatini ve derse katılımını öğrencinin yüz ifadelerine dayalı olarak tespit etmişlerdir. Çin'de ise öğrencilerin yüz ifadelerini tespit ederek öğrenci dikkatini belirleyen bir sistem geliştirilmiş ve bu sistem pilot uygulama olarak bazı okullarda kullanıma alınmıştır (Connor, 2018).

Afroze ve Hoque (2016) tarafından yapılan çalışmada görüntü işleme teknikleri kullanılarak kişinin kafa yönü bulunmak suretiyle kişiye ait dikkat odağı tespit edilmeye çalışılmıştır. Bunun için kameralardan birisi kullanıcıya dönük, diğeri ise kullanıcının baktığı yere dönük olacak şekilde iki kameralı bir sistem tasarlanmıştır. OpenCV ve Stasm kütüphaneleri kullanılarak yazılmış olan program vasıtasıyla, kişiye bakan kameradan alınan görüntülerden kişinin dik eksenindeki kafa yönü (sağ, sol, düz), diğer kameradan alınan görüntülerden de ortamdaki objeler tespit edilerek kişinin hangi objeye dikkatini verdiği

tahmin edilmeye çalışılmıştır. Kafa yönünün tespitinde nokta temelli geometrik analiz yöntemi kullanılmış, göz ile burun ucu arasındaki Öklid uzaklıklarından faydalanılmıştır. Objelerin tespitinde ise sobel kenar bulma algoritması ve morfolojik görüntü işleme tekniği kullanılmıştır. Bakış yönünün, objelerin ve hangi yöndeki objeye bakıldığıнын tespitinde tatmin edici sonuçlar alındığı belirtilerek deney sonuçları paylaşılmıştır.

Dikkat dağınıklığının tespiti ile ilgili çalışmaların önemli bir bölümü de sürüş güvenliği kapsamında yer almaktadır. Sürücü dikkatsizliğinin tespit edilmesi, görüntü işleme alanındaki popüler araştırma konularından birisidir. Yapılan araştırmalar trafik kazalarının %78'inin sürücü dikkatsizliğinden kaynaklandığını ortaya koymaktadır (Neale, Dingus, Klauer, Sudweeks ve Goodman, 2005). Sürücü dikkatsizliğinin tespiti ile ilgili çalışmalarda kullanılan yöntemlerin başında ise göz ve kafa hareketlerinin/yönünün takip edilmesi gelmektedir. Literatürde bu konuyla ilgili yapılmış pek çok çalışma olmakla birlikte bu alanda çığır açan isimlerin başında Alexander Zelinsky gelmektedir. Zelinsky 2000 yılında Avustralya Ulusal Üniversitesi ve Volvo Teknoloji'nin desteğiyle Seeing Machines isimli araştırma şirketini kurmuş; 2001 yılında ise 1996'ların sonlarında geliştirmeye başladığı faceLAB isimli sürücü destek sisteminin ilk ticari sürümünü yayınlamıştır (Apostoloff ve Zelinsky, 2003). Bu uygulama sürücüyü izleyen iki kamera yardımıyla kafa ve göz takibi yapmakta ve sürücünün nereye baktığını, sürücünün göz kapalılığını ve göz kırpmalarını tespit etmektedir. Tüm bu girdiler analiz edilerek sürücünün dikkat durumu hesaplanmakta ve dikkat durumu kritik seviyelere gelmeden sürücüyü uyarı verilmektedir (Zelinsky, 2006). Bu teknoloji otomotiv endüstrisinde yankı uyandırmış ve sürücü güvenliği uygulamaları alanında önemli bir ilerleme olarak kabul edilmiştir. Volvo, DaimlerChrysler, Robert Bosch ve Delphi gibi otomotiv devleri bu teknolojiyi satın almışlardır (Morais, 2004).

Kişinin görsel dikkat odağının tespitinde kullanılan parametrelerden bir diğeri de göz yönüdür. Göz yönünün tespiti de görüntü işleme teknikleriyle veya kafaya takılan göz izleme cihazlarıyla yapılabilmektedir.

Göz yönü bulmada görüntü işleme tekniklerinin kullanılması maliyet bakımından diğer yöntemlerden avantajlıdır. Fakat bu yöntemden başarılı sonuç alabilmek göz bölgesinin yeterli çözünürlükte görüntülerinin olmasına ve irisin ayırt edilebilir olmasına bağlıdır. Göz izleme cihazları ise pahalı olmalarına rağmen yüksek duyarlılıkta sonuçlar verdikleri için

tercih edilmektedir. Bu alanda Google Scholar’da listelenen akademik çalışmalarda en çok kullanılan cihazların:

1. Tobii (9230 yayın)
2. SMI(SensoMotoric Instruments) (6040 yayın)
3. SR research-EyeLink (5530 yayın)
4. ISCAN (2650 yayın)
5. LC Technologies (1130 yayın)
6. EyeTech (441 yayın)
7. The Eye Tribe (227 yayın)
8. Ergoneers (185 yayın)
9. Smart Eye (149 yayın)
10. Mirametrix (92 yayın)
11. Pupil Labs (80 yayın)
12. Gazepoint (39 yayın)

firmaları tarafından üretilmiş olan cihazlar olduğu bildirilmektedir (Farnsworth, 2017).

Göz takibi uygulamaları, nöroloji, psikoloji, endüstriyel mühendislik, pazarlama/reklamcılık ve bilgisayar bilimleri gibi alanlarda yaygın olarak kullanılmaktadır. Duchowski (2002) tarafından yapılan literatür tarama araştırmasında bu alanlarda yapılmış olan çalışmalar listelenmekte ve özetlenmektedir.

Ömur ve Görgülü Aydoğdu (2017) da göz izleme arařtırmaları ile ilgili kapsamlı bir literatür derlemesi yapmışlardır. Göz izlemenin ne olduğunu açıkladıktan sonra, göz izlemenin tarihsel gelişimine, göz izleme teknolojisine ve kullanım alanlarına değinmişlerdir. Çağıltay (2011)'dan aktardıklarına göre, “göz hareketleri verileri, insanların nereye dikkat ettiği, hangi bilgiyi göz ardı ettikleri, en fazla nelerden rahatsız oldukları gibi konular hakkında bilgi vermektedir”.

Cöster ve Ohlsson (2015) OpenCV kütüphanesini ve Viola-Jones yüz tespit algoritmasını kullanarak insan dikkatini ölçmeyi hedefleyen bir çalışma yapmışlardır. Bu kapsamda geliştirilen uygulama tarafından içerisinde yüz tespit edilen ve tespit edilemeyen fotoğraflar tespit edilmiştir. Bir kişi bir objeye doğru bakıyorsa kişinin dikkatinin o objeye odaklı olduğu kabul edilmiş ve içerisinde yüz tespit edilemeyen, yani kafa tespiti yapılamayacak kadar başka yöne bakmakta olan fotoğraflar dikkatsiz olarak varsayılmıştır. Yüz tespit

algoritması, ders sırasında farklı yönere bakmakta olan ve her fotoğrafta bir öğrenci olacak şekilde bir veri seti ile test edilmiştir. Fotoğraflarda kişilerin ortalı olmasına ve arka planda başka nesnelere olmamasına dikkat edilmiştir. Daha kapsamlı bir değerlendirme yapabilmek adına bu uygulama OpenCV kütüphanesi ile birlikte gelmekte olan 3 ayrı önceden eğitilmiş yüz tespit dosyası kullanılarak gerçekleştirilmiştir. Veri setindeki fotoğraflar, belirtilen kabuller doğrultusunda dikkatli ve dikkatsiz olarak gruplandırılmıştır. Bu sınıflandırma sonucunda 108 fotoğraf dikkatli, 127 fotoğraf ise dikkatsiz olarak belirlenmiştir. Kontrollü ortamda gerçekleştirilen bu deneyin başarılı sonuç verdiği ifade edilmiştir.

Ulusoy (2015) tarafından yapılan yüksek lisans tez çalışmasında ise dikkat dağınıklığının tespitinde kafaya takılan ve beyin sinyallerini okuyup kullanıcının dikkat ve duygu durumunu gerçek zamanlı olarak tespit edebilen NeuroSky MindWave Mobile cihazı kullanılmıştır. Test amaçlı olarak KonsantreOL isimli bir oyun geliştirilmiş, NeuroSky MindWave Mobile başlığı yardımıyla da oyunu oynayan öğrencilerin anlık dikkat değerleri ölçülmüştür. Oyun sırasında öğrenciye her 10 saniyede bir "Bilgisayar Programlama Dilleri 1" dersi ile ilgili sorular sorulmuş ve bu sorulara verilen yanıtlar kaydedilmiştir. Oyundaki amacın dikkat seviyesini belirlenen değerin üzerinde tutmak ve bu değerin devamlılığını sağlamak olduğu belirtilmiştir. Bunu yapabilmek için de dikkat değeri ölçülürken kaşları çatmanın ve yoğunlaşmanın gerektiği ifade edilmiştir.

Bu çalışma ilgili dersi almış olan Sakarya Üniversitesi Mekatronik Mühendisliği'nden 10 ve Karabük Üniversitesi Bilgisayar Mühendisliği'nden 44 öğrenci üzerinde uygulanmıştır. Geliştirilen oyun yardımıyla öğrencilerin dikkat seviyeleri ve dersle ilgili başarılarının ölçümü yapılmış ve dikkat dağınıklığının başarıya olan etkisi araştırılmıştır. Değerlendirme aşamasında öğrencilerin oyun sonunda elde ettiği dikkat başarıları sonuçları ve öğrencilerin "Bilgisayar Programlama Dilleri 1" dersi ara sınavından aldığı notlar karşılaştırmalı olarak incelenerek belirlenen ölçütlere göre dikkat seviyesinin ders başarısına etkisi incelenmiştir. Elde edilen verilere göre 44 kişilik öğrenci grubundan 24 kişide dikkatin ders başarısına etkisi görülürken 20 kişide dikkatin ders başarısını etkilemediği görülmüştür. Bu sonuçlara göre, dikkatin ders başarısına etkisi görülmekte ise de benzer çalışmaların daha çok yapılmasına ihtiyaç olduğu belirtilmiştir.

Esas (2017) tarafından yapılan “EOG testi ile dikkat eksikliđinin tanısı” konulu yüksek lisans tezinde dikkat dađınıklığına tespitinde Elektrokülogram yöntemi kullanılmıştır. EOG (Elektrokülogram) göz hareketlerinin oluşturduđu elektriksel aktivitenin göz bölgesinden yüzey elektrotları ile ölçümü olarak ifade edilmektedir. Tez çalışmasında EOG sinyallerinin elde edilebilmesi için elektronik devre tasarımı gerçekleştirilmiştir. Elektronik devre tarafından alınan sinyaller mikrodenetleyici vasıtasıyla sayısal hâle getirilerek USB portu üzerinden bilgisayara aktarılmıştır. Bu sinyallerinin kaydedilip izlenebilmesi için de bir ara yüz yazılımı gerçekleştirilmiştir. Sinyallerin elde edilmesi sırasında göz hareketlerinin analizinin sağlanması için bir dikkat test yazılımı geliştirilmiş ve test üzerinde referans yönler belirlenmiş olup gözün belirlenen sürede belirlenen yönler göre kaç kez hareket ettiđi tespit edilmiştir. Test eşliğinde EOG sinyalleri Dikkat Eksikliđi Hiperaktivite Bozukluđu (DEHB) tanısı konulmuş 8-13 yaş arası 8 hastada ve aynı yaş grubuna ait 8 sağlıklı bireyden kaydedilmiştir. Daha sonra bu sinyaller bilgisayar ortamında analiz edilmiştir. Yapılan çalışmalar sonucu sağlıklı bireylerde test performans genel başarı oranı %93,11 tespit edilirken hasta bireylerde % 85,14 olarak tespit edilmiştir. Test sonucu her bir periyottan elde edilen hareket sayısı ortalamasının doğruluk yüzdesine t-Testi uygulanmış olup, sağlıklılarda anlamlılık düzeyi (p), 0,11 hastalarda ise 0,049 olarak elde edilmiştir. Bu sonuçlara göre sağlıklı bireylerin göz hareketleri sayısının istatistiksel olarak anlamlı olduđu ve referans deđer ile uyumlu olduđu, hastalarda ise anlamlılık düzeyinin düşük olduđu görülmüştür.

2.2. Yüz Tanıma ile İlgili Çalışmalar

Yüz tanıma biyometrik tanıma kapsamında yer alan ve en çok kullanılan tekniklerden birisidir. Yüz tanıma ile ilgili çalışmalara geçmeden önce biyometriden ve biyometri ile ilgili çalışmalardan kısaca bahsetmek faydalı olacaktır.

Canlılara ait her türlü gözlemlerde uygulanan istatistik biçimi olarak tanımlanan biyometri kelimesi, Yunanca'da hayat anlamına gelen "bios" ve ölçü anlamına gelen "metron" kelimelerinin birleşmesinden oluşmaktadır (Karol, 1963).

Bireyi eşsiz kılan biyometrik özellikler yaşamın başlangıcından sonuna kadar deđiştirilemez bir yapıya ve yüksek güvenlik hassasiyetine sahiptir. Yüksek düzeyde doğruluk gerektiren

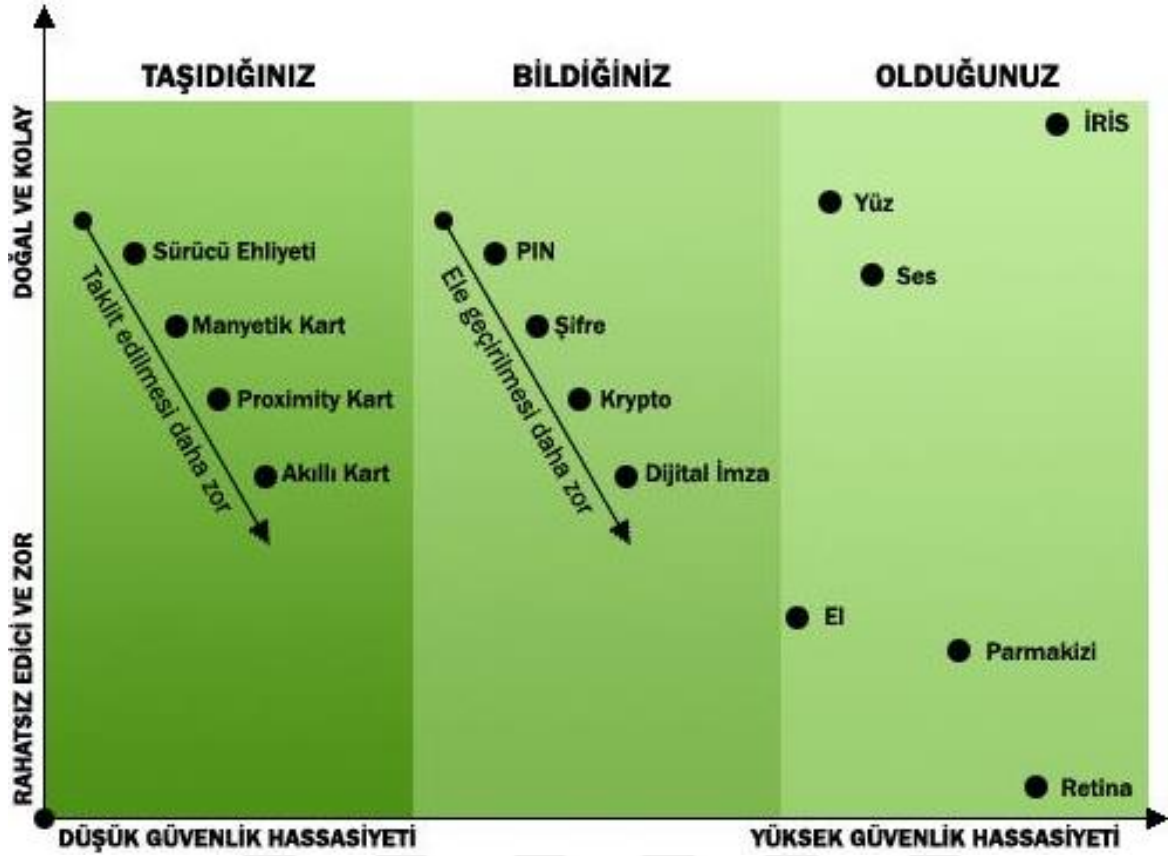
uygulamalara duyulan ihtiyaçtan dolayı da biyometri üzerine yapılan çalışmalar her geçen gün artmaktadır (Uğur, 2018).

Juniper Research isimli firma tarafından 08.05.2017'de yayınlanan bir araştırma raporunda, 2016 yılında dünya genelinde 600 milyon adet biyometrik doğrulamalı mobil ödeme gerçekleştirirken, 2017 yılında bu sayının 2 milyara ulaşacağı tahmini edilmiştir (Smith, 2017).

Yine aynı firma tarafından 24.07.2018'de yayınlanan araştırma raporunda ise, 2018 yılında biyometri doğrulaması aracılığıyla yapılan mobil ödeme işlemlerinin toplam değerinin 124 milyar\$ olması beklenirken, 2023 yılında bu rakamın 2 trilyon\$a çıkacağı yani yaklaşık 17'ye katlayacağı tahmin edilmektedir (Smith, 2018).

Biyometriye dayalı teknikler, son yıllarda kişilerin tanınmasında en umut verici seçenek olarak ön plana çıkmıştır. Kişilerin fiziksel ve sanal ortamlarda yetkilendirilmesinde ve kimlik doğrulamasında kullanılan şifrelerin, PIN'lerin, akıllı kartların aksine biyometriye dayalı teknikler kişilerin kimliklerinin teşhisinde onların bu biricik özelliklerini kullanır. Şifreler, PIN'ler çalınabilir, tahmin edilebilir ya da hatırlaması zor olabilir; kartlar bozulabilir, kaybedilebilir, unutulabilir ya da kopyalanabilir. Fakat kişinin biyometrik özellikleri unutulamaz, kaybedilemez, çalınamaz ve taklit edilemez (Jafri ve Arabnia, 2009).

Geleneksel yöntemlere ve biyometriye dayalı kimlik doğrulaması ile ilgili karşılaştırma Şekil 2.2'de görülmektedir.



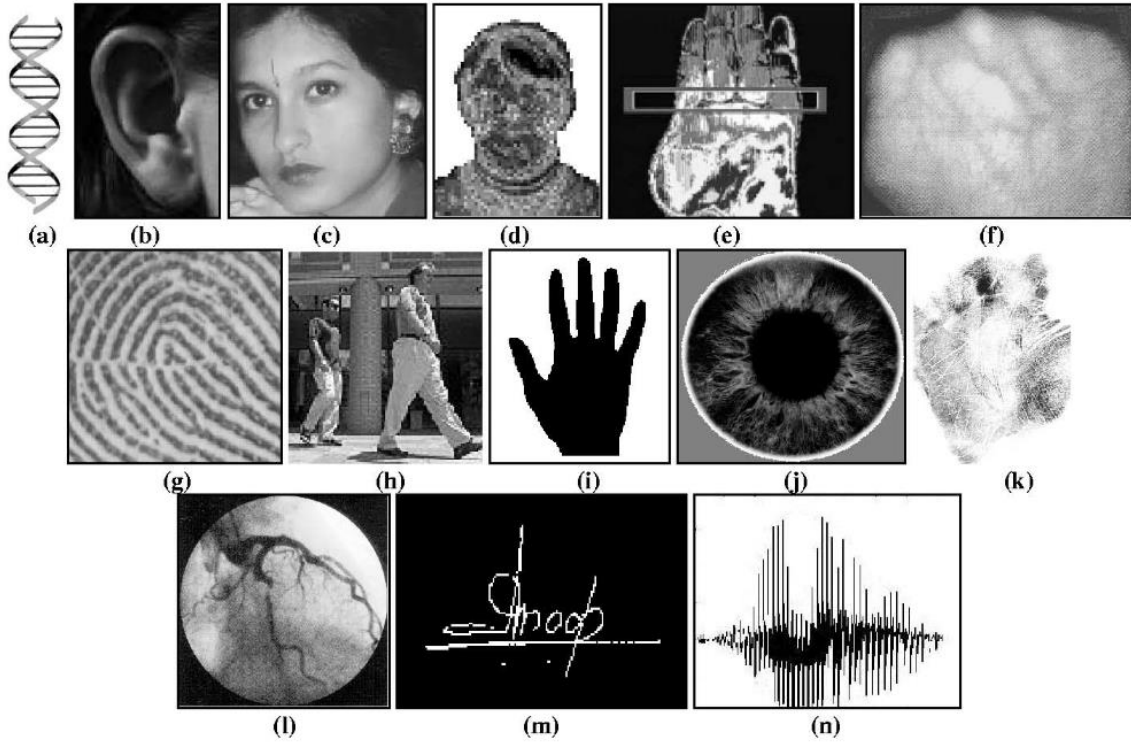
Şekil 2.2. Geleneksel yöntemlerle biyometriye dayalı kimlik doğrulamasının karşılaştırılması (Uğur, 2018)

Biyometriye dayalı tanıma teknikleri kişilerin onlara özgü olan fizyolojik ve davranışsal özelliklerinden yararlanmaktadır (Jain, Ross ve Prabhakar, 2004).

Fizyolojik özelliklere dayalı tanımda yüz, kulak, parmak izi, el damarları, avuç içi, el geometrisi, iris, retina vb. özellikler sıkça kullanılırken; davranışsal özelliklere dayalı tanımda ise yürüyüş şekli, imza, tuşlara basma şekli gibi karakteristikler kullanılmaktadır.

Ses ise bazı yayınlarda fizyolojik karakteristikler arasında sınıflandırılmakta iken (Jafri ve Arabnia, 2009), bazı yayınlarda davranışsal karakteristikler arasına dâhil edilmiştir (Bhattacharyya, Ranjan, Alisherov ve Choi, 2009; Yampolskiy ve Govindaraju, 2008). Jain ve diğerleri (2004) ise sesin hem fizyolojik hem de davranışsal bir karakteristik olduğunu ifade etmiştir.

Bahsi geçen biyometrik karakteristiklerden bazıları Şekil 2.3’de örneklenmiştir.



Şekil 2.3. Biyometrik karakteristik örnekleri: (a) DNA, (b) kulak, (c) yüz, (d) yüz termogramı, (e) el termogramı, (f) el damarları, (g) parmak izi, (h) yürüyüş şekli, (i) el geometrisi, (j) iris, (k) avuç içi, (l) retina, (m) imza, (n) ses (Jain ve diğerleri, 2004).

Biyometrik doğrulama, gelişmekte olan ve yaygınlaşan bir alan olmasının yanı sıra kişilerin mahremiyetlerinin ihlal edilmesine yol açtığı iddialarıyla da karşı karşıya kalmaktadır. Mahremiyet ihlalleri ile ilgili iddiaların odağında ise son yıllarda pek çok alanda yaygın olarak kullanılan ve kameralarla kişilerin izlenmesini gerektiren yüz tanıma sistemleri gelmektedir.

Teknolojideki hızlı ilerlemeler ve terör tehdidi gibi sebepler tanınacak kişi ile bir etkileşim gerektirmeyen yüz tanıma teknolojisinin yaygınlaşmasında tetikleyici rol oynamıştır. Bununla beraber, yüz tanıma konusuna artan ilgi ve bu teknolojinin yaygın olarak kullanılmaya başlanması mahremiyet odaklı tartışmaları da beraberinde getirmiştir.

Özellikle Amerika'da yaşanan 11 Eylül 2001 tarihli terör saldırısı sonrasında bireylerin kamerayla takip edilmesi, yüz görüntülerinden hareketle kişilerin tanınması, aranan kişilerin, şüphelilerin önceden tespit edilmesi konuları ön plana çıkmıştır. Bu saldırının hemen sonrasında yapılan ve sonraki yıllarda da tekrarlanan bir anket çalışmasında sorulan sorulardan ikisi halkın çoğunluğunun yüz tanıma teknolojisine ve halka açık yerlerde kamera

ile gözetleme yapılmasına destek verdiğini, karşı olanların yüzdesinin ise ihmal edilmeyecek düzeyde olduğunu göstermektedir (Çizelge 2.1).

Çizelge 2.1. 2001-2005 yılları arasında yapılmış olan Harris-Westin anket sonuçları (Westin, 2006)¹.

Anket Sorusu	Tarih	Destekliyorum	Karşıyım	Kararsızım
<i>Halka açık etkinliklerde ve çeşitli bölgelerde şüpheli kişilerin, teröristlerin tespit edilebilmesi amacıyla yüz tanıma teknolojisinin kullanılması</i>	Şubat 2004	%80	%17	%3
	Şubat 2003	%77	%20	%3
	Mart 2002	%81	%17	%2
	Eylül 2001	%86	%11	%2
<i>Halka açık yerlerde, sokak ve caddelerde kamera ile genişletilmiş bir gözetleme yapılması</i>	Haziran 2005	%59	%40	%1
	Şubat 2004	%61	%37	%2
	Şubat 2003	%61	%37	%1
	Mart 2002	%58	%40	%2
	Eylül 2001	%63	%35	%2

Anlaşıldığı üzere, yüz tanıma teknolojisi özellikle güvenlik kaygısı ile desteklenmekte ve kullanılmakta iken, öte yandan bu teknolojinin halka açık ortamlarda kullanımına mahremiyet ihlali açısından karşı çıkanlar bulunmaktadır.

Bowyer (2004)'in yapmış olduğu çalışma, yüz tanıma teknolojisini güvenlik-mahremiyet kavramları bağlamında değerlendirmekte, bu teknolojinin kullanımının sağlayacağı faydalarla birlikte bu teknolojinin kullanımına yönelik eleştirilere de yer vermektedir. Kamera ile kişilerin izlenmesinin Büyük Birader (Big Brother)² benzeri korkulara sebep olduğu belirtilmektedir. Bu teknolojinin kullanımına karşı çıkanların başında Amerikan Sivil Özgürlükler Birliği (ACLU, American Civil Liberties Union)'nin ve Elektronik Sınırlar Vakfı (EEF, Electronic Frontiers Foundation)'nin geldiği belirtilmekte ve Amerikan Sivil Özgürlükler Birliği'nin bazı çelişkili eleştirileri de bu makalede örneklenmektedir.

Bu ön bilgilerden sonra, yüz tanımanın ne olduğuna ve bu alandaki ilk çalışmalara değinebiliriz.

¹ Anketteki sorulara verilen cevaplar yüzdelik olarak gösterilmiştir. Yüzdelik sonuçların toplamının yüz olmamasının yuvarlamadan kaynaklandığı belirtilmiştir.

² Orwell'in 1984 isimli kitabında Okyanusya ülkesinin merkezi otoritesini temsil eden kavram. Bu ülkede toplumdaki tüm bireyler *Büyük Birader* tarafından sürekli takip edilmekte ve izlenmektedir.

Yüz tanıma, dijital bir görüntüdeki ya da video karesindeki kişilerin kimliklerinin tespit edilmesini ya da doğrulanmasını mümkün kılan biyometriye dayalı bir tanıma tekniğidir. Parmak izi tanıma kadar yaygın kullanılsa da biyometriye dayalı tanıma teknikleri içinde en çok kullanılan tekniklerden birisidir. Parmak izine dayalı doğrulama güvenilir doğruluk oranlarına sahipken yüz tanıma için bu oranlar çeşitli sebeplerle (aydınlatma, görüntü kalitesi, sakal, bıyık, gözlük vs.) daha düşük kalmaktadır. Fakat parmak izi tanımada, bir parmak izi okuyucu cihaza ve kişinin bu cihaza parmağını okutmasına ihtiyaç duyulurken yüz tanımda sadece kişilerin görüntülerini çekmek için bir kamera yeterli olmaktadır. Bu da yüz tanımanın son zamanlarda yaygınlaşmasının en önemli nedenlerindedir.

Bir yüz tanıma sisteminde temel olarak yapılan iş sadedir. Bir kamera vasıtasıyla belli bir alan (bir havalimanındaki inen yolcuların olduğu bölüm vb.) izlenir. Bir kişinin devamlı olarak kameradan alınan görüntüyü izlemesi yerine, geliştirilen yüz tanıma yazılımı vasıtasıyla görüntünün taranması ve istenilen kişi tespit edildiğinde de uyarı verilmesi amaçlanır. Bu da tespit edilmesi istenilen kişilerin resimlerinin yüz tanıma yazılımının veritabanına kaydedilmesini gerektirmektedir.

Bilgisayar destekli yüz tanımda ilk adım video görüntüsündeki yüzün tespit edilmesidir (face detection). Yüz tespit edildikten sonra video görüntüsünden elde edilen yüz kırplır, aydınlatma, resim boyutu gibi faktörler göz önünde bulundurularak normalize edilir (preprocessing). Yüzdeki öznitelikler çıkarılır (feature extraction) ve bu öznitelikler kullanılarak eğitim (training) gerçekleştirilir. Veritabanındaki öznitelikler arasından en çok uyumlu olan resim ile eşleştirme yapılarak (classification) yüz tanıma gerçekleştirilmiş olur (face recognition).

Bilgisayar destekli yüz tanıma konusunda yapılan çalışmaların öncüleri Woodrow Wilson (Woody) Bledsoe, Helen Chan Wolf ve Charles Bisson'dur. 1964 ve 1965 yıllarında, Bledsoe, Chan ve Bisson ile birlikte bilgisayarı insan yüzlerini tanımda kullanma üzerine çalışmalar yaptı (Bledsoe, 1966a; Bledsoe, 1966b; Bledsoe ve Chan, 1965).

Bledsoe yaptığı çalışmalardan gurur duyuyor olsa da ona sponsor olan bir istihbarat teşkilatının sonuçların halka açık olarak paylaşılmasına çok fazla izin vermemesinden ötürü çalışmalarının çok küçük bir bölümü yayınlanmıştır. Üzerinde çalıştığı problem, bir kitap

dolusu resimden oluşan geniş bir veritabanı içerisinde verilen bir fotoğrafa benzeyen ve içlerinden biri verilen fotoğraf ile eşleşen küçük bir resim kümesini seçmekti. Yöntemin başarısının ise cevap listesinin veritabanındaki toplam resim sayısına oranı cinsinden ölçülebileceği belirtilmiştir (Ballantyne, Boyer ve Hines, 1996).

Bledsoe (1966a), yüz tanıma problemindeki zorluklara örnek olarak kafa yönü, ışıklandırma yoğunluğu ve açısı, yüz ifadeleri ve yaş gibi parametrelerdeki büyük değişkenlikleri vermektedir. Yapmış olduğu çalışmada, fotoğraflardaki öznitelik kümesine (gözbebeğinin merkezi, gözün iç köşesi, gözün dış köşesi, saç-alın çizgisinin orta noktası vb.) ilişkin koordinatlar bir insan tarafından çıkarılmakta ve bu koordinatlar bilgisayar tarafından kullanılmaktadır. Bundan ötürü projesini man-machine olarak adlandırmıştır. Çıkarılan bu koordinatlardan 20 farklı uzaklık (dudak genişliği, gözlerin genişliği, göz bebekleri arası uzaklık gibi) hesaplanmıştır. Bu işi yapan kişilerin saatte 40 fotoğrafı işleyebildikleri belirtilmiştir. Veritabanını oluştururken, bir fotoğrafa ilişkin uzaklıklar ile fotoğraftaki kişinin ismi ilişkilendirilerek bilgisayara kaydedilmiştir. Tanıma aşamasında ise, sorgulanan fotoğrafa ilişkin uzaklıklar veritabanındaki her bir fotoğraftaki uzaklıklar ile kıyaslanmış en uyumlu olan kayıtlar bulunmuştur.

Teknolojinin ilerlemesi ile birlikte son yıllarda yüz tanıma konusunda da önemli ilerlemeler kaydedilmiş ve bu alanda çığır açan bazı yöntemler ortaya çıkmıştır. Şimdi bu yöntemlerden ve çalışmalardan başlıcalarına değinelim.

Özyüzler (Eigenfaces) yöntemi, ilk defa Sirovich ve Kirby (1987) tarafından yüzü etkin bir biçimde ifade edebilmek için kullanılmıştır. Bu yöntem istatistiksel bir yaklaşım olan Temel Bileşenler Analizi'ne (Principal Component Analysis, PCA) ya da Karhunen–Loève genişlemesine dayanmaktadır. Bu yöntemin amacı, resimlerdeki çeşitliliği en çok etkileyen temel bileşenlerin çıkarılmasıdır. Bu sayede yüz uzayındaki resimler daha az boyutlu bir alt uzayda ifade edilebilmektedir. Bütüncül (holistic) bir yaklaşıma dayanmaktadır.

Özyüzler yöntemi yüz uzayının temel bileşenlerini ve yüz vektörlerinin temel bileşenler üzerine izdüşümünü kullanır. Bu yöntem boyut indirgemeye dayalı olduğundan tanıma ve öğrenme işlemleri hızlıdır. Bu yöntemde, yüz uzayında bulunan resimlerin değerleri arasındaki varyanslardan yararlanır. Örneğin bir x değeri y ile bağlantılı ise ve x 'in değerini

bilmek y'nin deęerini ifade etmede yeterli ise (x,y) iki boyutlu uzayı yerine tek boyutlu x deęerlerini tutmak yeterli olacaktır (Kekül, Bircan ve Arslan, 2018).

Turk ve Pentland (1991) ise yapmış oldukları çalışmada bu yöntemi yüz tanıma için daha kararlı bir hâle getirmişlerdir. Bu sayede bu yöntemin kullanımını da ivme kazanmıştır.

Fisherfaces yöntemi özyüzler yöntemine alternatif olarak ortaya çıkmış bütüncül (holistic) yaklaşıma dayalı bir yöntemdir (Belhumeur, Hespanha ve Kriegman, 1997). Özyüzler yöntemine benzer olarak Fisherfaces yönteminde de boyut indirgeme önemli bir adımdır. Bu amaçla özyüzler yönteminde kullanılan TBA yerine FisherFaces yönteminde Fisher (1936)'ın formüle ettiği Doğrusal Ayırtaç Analizi (Linear Discriminant Analysis) kullanılır. Bu yöntemi özyüzler yönteminden farklı kılan en önemli özellik ise sınıflandırmada kullanılacak olan resimlerin etiketlerinin DAA'da önem taşıması denilebilir. Yani FisherFaces yöntemine dayalı bir öğrenme gözetimli (supervised) bir öğrenme iken TBA'ya dayalı bir öğrenme gözetimsiz (unsupervised) bir öğrenmedir.

Örüntü tanıma uygulamalarında yaygın olarak kullanılan bir diğer yöntem de Yerel İkili Örüntüler (Local Binary Patterns) yöntemidir. Bu yöntem ilk olarak Ojala ve arkadaşları tarafından öne sürülmüştür (Ojala, Pietikäinen ve Harwood, 1994, 1996). Bu yöntem yerel komşuluğa dayanarak özniteliklerin istatistiksel olarak çıkarımını sağlar, bütüncül (holistic) bir yöntem değildir. Işık deęişikliklerinden çok etkilenmemesi de raębet görme nedenlerinden birisidir. Ahonen ve arkadaşlarının yapmış olduęu çalışma ile bu yöntemin yüz tanıma uygulamalarındaki başarısı ortaya konmuştur (Ahonen, Hadid ve Pietikäinen, 2006).

Yüz tanıma konusunda son yıllarda yaygın olarak kullanılmaya başlanan bir diğer yöntem ise derin öğrenmedir. Derin öğrenme, makine öğrenme algoritmaları kullanılmak suretiyle doğrusal olmayan dönüşümler vasıtasıyla tanınacak olan veriyi en iyi temsil edecek özellikleri bulmayı amaçlayan bir yaklaşımdır. Derin öğrenme isimlendirmesindeki “derin” kavramı klasik yapay sinir ağlarına göre daha çok katmanlı bir yapıya dayalı bir öğrenmeyi belirtmektedir.

Temelinde yapay sinir ağları olan derin öğrenme insan beyninin çalışmasını yapay sinir ağlarına göre daha iyi modelleyen yöntemler önerir. Derin öğrenmede kullanılan

algoritmalarından birçoğu Geoffrey Hinton ve Toronto Üniversitesi'ndeki araştırmacıların yaptıkları çalışmalara dayanmaktadır. 1980'lerde öne sürülen bu algoritmalar yoğun matris işlemleri gerektirdiğinden ve büyük verilerin işlenmesinde yüksek işlemci gücüne ihtiyaç duyduğundan o yıllarda yaygın bir uygulama alanına kavuşamamıştır.

1980'lerin sonlarında uygulamaları yapılmış olan Geriye Yayıma Algoritması (Rumelhart, Hinton ve Williams, 1986; LeCun, 1988) bilim çevrelerinde büyük yankı uyandırmamışsa da Kanada İleri Araştırmalar Enstitüsü'nün dikkatini çekmiştir. Belki de bu sayede enstitüye bağlı Montreal ve Toronto üniversitelerindeki araştırma grupları derin öğrenme alanında referans verilen gruplar olarak öne çıkmıştır (Hacıoğlu, 2014).

Özellikle görüntü işleme alanında derin öğrenmenin özelleşmiş bir mimarisi olan Evrimsel Sinir Ağları (Convolutional Neural Networks) başarılı sonuçlar vermekte ve son zamanlarda daha da yaygın olarak kullanılmaktadır.

Bu mimaride öznitelik çıkarımı, boyut azaltma, eğitim ve sınıflandırma gibi işleri gerçekleştirebilmek amacıyla çeşitli katmanlar bulunur. Bir ESA'da bulunan katmanların başlıcaları konvolüsyon katmanı (convolution layer), düzleştirilmiş doğrusal birim katmanı (rectified linear units layer, ReLu) ve havuzlama (pooling layer) katmanıdır. Son bölümde ise sınıflandırma işlemini gerçekleştiren ve standart bir yapay sinir ağı olan tam bağlı katman (fully-connected layer) bulunur.

ESA'da giriş verileri alındıktan sonra katman katman işlemler yapılarak eğitim gerçekleştirilir. Eğitim sonunda beklenen sonuç ile ESA'nın ürettiği sonuçlar kıyaslanır. Aradaki fark bulunarak bu hata payının bütün ağırlıklara aktarılması için geriye yayılım algoritması kullanılır. Her bir iterasyonla ağırlıkların güncellenmesi yapılır ve hatanın azaltılması sağlanır (İnik ve Ülker, 2017).

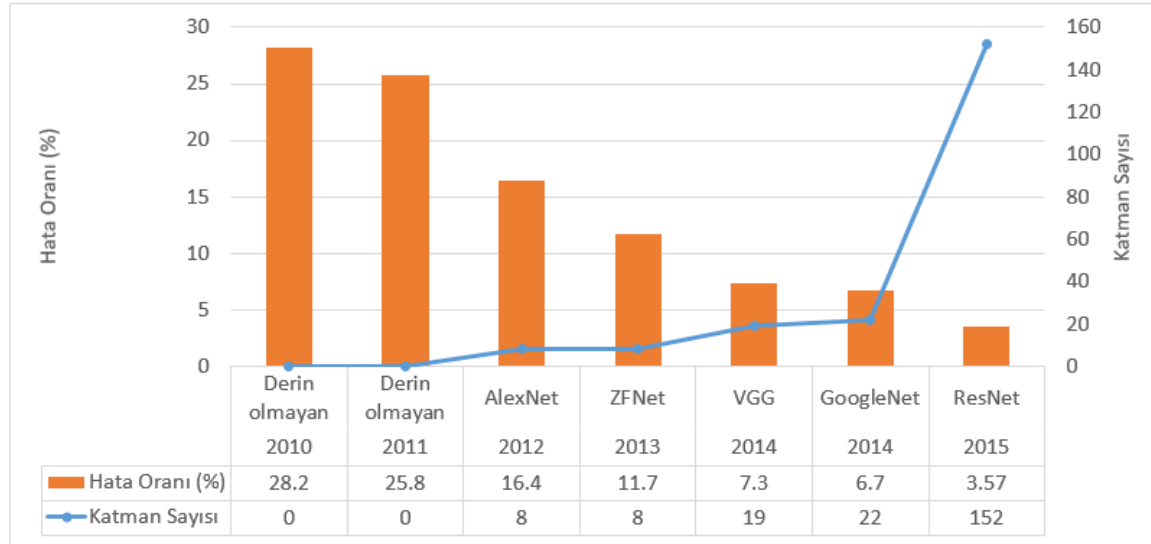
Derin öğrenmenin temellerini oluşturan ESA'nın bilim dünyasında büyük yankı uyandırması ve popülerlik kazanması, 2012 yılında düzenlenmiş olan ImageNET Büyük Ölçekli Görsel Tanıma Yarışması'nı (ILSVRC - ImageNet Large Scale Visual Recognition Challenge 2012) Alex Krizhevsky, Ilya Sutskever ve Geoffrey E. Hinton'dan oluşan takımın bu yöntemi kullanarak dikkat çeken bir başarı elde ederek kazanmasıyla başlamıştır diyebiliriz.

Bu yarışmadaki sonuçların başarısını ölçmede kullanılan Top-5 hata oranı Krizhevsky ve diğerlerinin AlexNet isimli çalışmasıyla %16,4'e düşürülmüştür (Krizhevsky, Sutskever ve Hinton, 2017). İkinci sıradaki çalışmanın hata oranının yaklaşık %26 olması ve aradaki yaklaşık %10'luk fark bilgisayarla görme camiasında şok edici bir etki yaratmıştır.

ILSVRC 2012'de Krizhevsky ve diğerlerinin derin öğrenme yöntemi ile birinci olmasının ardından, derin öğrenmenin ve ESA'nın görüntü sınıflandırmadaki başarısı görülmüş, ertesi yılki yarışmada da birçok takım derin öğrenme tekniğini kullanarak başarılı sonuçlar elde etmiştir. Öyle ki, ilk 20'ye girenlerin tamamı da derin öğrenme tekniğini kullanmışlardır (Cengil ve Çınar, 2016).

ILSVRC 2013'ün kazananı ise %11.7 hata oranı ile New York Üniversitesi'nden Matthew Zeiler ve Rob Fergus'un derin öğrenmeye dayalı çalışması olmuştur (Zeiler ve Fergus, 2014).

Derin öğrenmedeki ilerlemeler top-5 hata oranını %3,6'lara kadar düşürmüştür (Şekil 2.4).



Şekil 2.4. ILSVRC yarışmasındaki yıllara göre Top-5 hata oranları ve kullanılan katman sayıları

3. MATERYAL ve YÖNTEM

Bu bölümde uygulama geliştirme aşamasında gereksinim duyduğumuz donanım, yazılım geliştirme ortamı, yazılım kütüphaneleri ile çalışmada kullanılan yöntemler açıklanmıştır.

3.1. Donanım

Belirtilen yazılımın geliştirilmesi ve test edilmesinde kullanılan bilgisayarın temel özellikleri Çizelge 3.1’de gösterilmiştir.

Çizelge 3.1. Kullanılan donanım.

İşlemci	Intel(R) Core(TM) i7-4770K CPU @ 3.50GHz
Ekran Kartı	Gigabyte NVIDIA GeForce GTX 980 4 GB
RAM	16 GB
Web kamerası	Logitech HD Webcam C525

3.1.1 Web kamerası

Web kamerası, üzerinde bulunan sensörler vasıtasıyla görüntüyü alıp bu görüntüyü dijital veriye dönüştüren aygıttır. Bu çalışmada saniyede 30 çerçeve görüntüyü 720p çözünürlükte yakalayabilen Logitech HD Webcam C525 web kamerası kullanılmıştır (Resim 3.1).



Resim 3.1. Logitech HD Webcam C525

3.2. Yazılım Geliştirme Ortamı ve Yazılım Kütüphaneleri

Yazılım geliştirme ortamı ve kullanılan yazılım kütüphaneleri Çizelge 3.2’de gösterilmiştir.

Çizelge 3.2. Yazılım geliştirme ortamı ve kullanılan yazılım kütüphaneleri.

İşletim Sistemi	Windows 10 Pro 64-bit
Yazılım Geliştirme Aracı	Qt Creator 4.2.1
Yazılım Kütüphaneleri	OpenCV 3.3.1, dlib 19.8

3.2.1. Qt Creator

Qt Creator, masaüstü, gömülü ve mobil platformlar için C++ ve QML uygulamaları geliştirmek için kullanılan platform bağımsız bir yazılım geliştirme ortamıdır. Qt Creator içerisinde kod düzenleyici ile birlikte gelmekte, yazılımın tasarımı, kodlanması, test edilmesi, dağıtılması ve bakımı için araçlar içermektedir. Basit ve kolay öğrenilen arayüzü, yazım vurgulaması, otomatik tamamlama, sürükle-bırak ile kullanıcı arayüzleri oluşturma, görsel hata ayıklama özellikleri ve proje geliştirmeyi destekleyen pek çok araç barındırması ile kullanıcıların tercih ettiği bir yazılım geliştirme ortamıdır. Bu çalışmada QT Creator 4.2.1 IDE’si ve c++ programlama dili kullanılarak bir uygulama geliştirilmiştir.

3.2.2. OpenCV

Bradski (2000) tarafından temelleri atılan OpenCV (Open Source Computer Vision Library), açık kaynak kodlu bir görüntü işleme ve makine öğrenmesi yazılım kütüphanesidir. OpenCV kütüphanesi, görüntü işleme uygulamalarına genel bir altyapı sağlamak ve ticari yazılımlardaki makine algılaması kullanımını özendirme amacıyla geliştirilmiştir. BSD lisanslı bir ürün olan OpenCV ticari kuruluşların kodu kullanıp değiştirmelerini kolaylaştırmaktadır (OpenCV, 2019).

OpenCV kütüphanesi 2500'den fazla optimize edilmiş algoritma içermektedir. Bu algoritmalar hem klasik hem de son teknoloji görüntü işleme ve makine öğrenmesi algoritmalarından oluşmaktadır. Bu algoritmalar yardımıyla yüz tespiti, yüz tanıma, nesnelerin tanımlanması, videolardaki insan davranışlarının sınıflandırılması, kamera hareketlerinin izlenmesi, hareketli nesnelerin takibi, nesnelerin 3 boyutlu modellerinin çıkarılması, bir görüntü veritabanındaki benzer görüntülerin tespiti, flaş kullanılarak çekilen fotoğraftaki kırmızı gözün giderilmesi, göz hareketlerinin izlenmesi gibi işlemler yapılabilmektedir. OpenCV kütüphanesi 47 binden fazla kullanıcıdan oluşan bir yazılım topluluğuna ve yaklaşık 18 milyondan fazla indirme sayısına sahiptir. Kütüphane şirketlerde, araştırma gruplarında ve devlet kurumlarında yaygın olarak kullanılmaktadır (OpenCV, 2019).

OpenCV kütüphanesi C++ ile yazılmış olup Python, Java ve MATLAB arayüzlerine de sahiptir ve Windows, Linux, Android ve Mac OS işletim sistemlerini desteklemektedir (OpenCV, 2019).

OpenCV 3.3.1 sürümünde bulunan önemli modüller şunlardır:

- **core:** Kütüphanenin temellerini oluşturan ve diğer modüller tarafından da kullanılan veri yapılarını içeren modüldür. Görüntüler üzerinde piksel seviyesinde işlem yapabilmek için bu modülün öğrenilmesine ihtiyaç vardır.
- **imgproc:** Görüntü işleme fonksiyonlarını içeren modüldür.
- **imgcodecs:** Görüntü dosyalarının diske yazılması ve diskten okunması ile ilgili modüldür.
- **videoio:** Video dosyalarının diske yazılması ve diskten okunması ile ilgili modüldür.

- **highgui:** Kütüphane içerisinde hazır olarak gelen grafik kullanıcı arayüzü modülüdür.
- **video:** Video analizi modülüdür. Video akışlarındaki hareketlerin tespiti, özniteliklerin takibi gibi işlerin yapılabilmesini sağlar.
- **calib3d:** Kameranın kalibre edilmesi ve 2 boyutlu görüntüden 3 boyutlu obje modelin elde edilmesi ile ilgili algoritmaların sağlandığı modüldür.
- **features2d:** Öznitelik noktaları tespit edicilerin ve ilgili araçların yer aldığı modüldür.
- **objdetect:** Objelerin, yüzün tespit edilmesi gibi ihtiyaçlar için gereken algoritmaları içeren modüldür.
- **dnn:** Son zamanlarda popüler olan derin öğrenme ile ilgili altyapının sağlandığı modüldür.
- **ml:** Akademik ve ticari amaçlarla yaygın olarak kullanılmakta olan ve sınıflandırma ve regresyon gibi problemlerin çözülebilmesine imkan tanıyan makine öğrenmesi algoritmalarının bulunduğu modüldür.
- **photo:** Fotoğrafların işlenmesi ile ilgili gelişmiş yöntemleri barındıran modüldür.
- **stitching:** Fotoğraflardan panorama yaratmak gibi işler için kullanılan algoritmaları barındıran modüldür.
- **cuda:** OpenCV algoritmalarının sistemin işlem gücünü daha efektif kullanabilmesine, ekran kartının işlemcisini (GPU) de kullanabilmesi için gerekli olan altyapıyı sağlayan modüldür.
- **face:** Ana modüller içerisinde olmamakla birlikte ekstra modüller arasında yer almaktadır. Yüz tanıma ile ilgili önemli algoritmaları barındıran modüldür.
- **fuzzy:** Ekstra modüller arasında yer almaktadır. Bulanık matematiğe dayalı görüntü işleme algoritmalarını barındıran modüldür (OpenCV, 2019).

3.2.3. Dlib

Dlib kütüphanesi -ya da yaratıcısı Davis King'in deyişiyle Dlib makine öğrenmesi araç kutusu (King, 2009)- gerçek dünyadaki problemleri çözmek için karmaşık yazılımların geliştirilmesini sağlayan araçlardan oluşan ve makine öğrenme algoritmalarını da içeren modern bir C++ yazılım kütüphanesidir (dlib, 2019).

Dlib kütüphanesi hem yazılım endüstrisinde hem de akademik camiada robotik, gömülü cihazlar, cep telefonları, yüksek performanslı hesaplama ortamları gibi geniş bir yelpazede kullanılmaktadır. Boost lisansına sahip olması Dlib'in her türlü yazılımda ücretsiz olarak kullanılmasına imkân vermektedir. Diğer çoğu açık kaynak kodlu projelerin ve kütüphanelerin aksine dlib kütüphanesi çok iyi bir dokümantasyona sahiptir ISO standart C++ ile yazılmış olup başka kütüphanelere bağımlılığı yoktur. MS Windows, Linux ve Mac OS X sistemlerinde sorunsuz olarak çalışmaktadır (dlib, 2019).

Dlib kütüphanesinde yer alan önemli algoritma grupları şunlardır:

- **Makine öğrenmesi algoritmaları:** Derin öğrenme, Destek Vektör Makineleri, Kümeleme algoritmaları (Chinese Whispers, Newman vb.), çok katmanlı algılayıcılar gibi çeşitli makine öğrenmesi algoritmalarını içermektedir.
- **Sayısal algoritmalar:** Bu kapsamda kullanılmak üzere hızlı çalışacak bir matris objesi ve bu matris objesi üzerinde gerçekleştirilebilecek çok sayıda lineer cebir ve matematik işlemleri (transpoz, tekil değer ayrışımı vs.) tanımlanmış, “big integer” ve “random number” veri tipleri oluşturulmuştur.
- **Görüntü işleme algoritmaları:** Yaygın kullanılan görüntü formatlarını diskten okuma ve diske yazma ile ilgili metotlar, çeşitli piksel tipleri arasında otomatik renk uzayı dönüşümü yöntemleri bu kategoride yer almaktadır. Ayrıca yaygın kullanılan görüntü işleme teknikleri (kenar bulma ve morfolojik işlemler gibi), SURF, HOG, FHOg gibi öznitelik çıkarımı algoritmaları, yüksek kalite yüz tanıma algoritmaları, görüntüdeki objelerin tespit edilmesi ile ilgili algoritmalar bu kategorideki başlıca öğelerdendir (dlib, 2019).

3.3. Kullanılan Yöntemler

Bu çalışma kapsamında geliştirilen uygulamada yüz bulma, Yönlendirilmiş Gradyanların Histogramı yöntemiyle; yüz tanıma, Yerel İkili Örüntüler yöntemiyle; dikkatli ve dikkatsiz olan öğrencilerin sınıflandırılması Destek Vektör Makineleri makine öğrenme algoritmasıyla; kafa yönü açılarının bulunması ise Perspective-n-Point probleminin çözülerek Euler açılarının hesaplanması yoluyla gerçekleştirilmektedir. Yüz bulma ve yüzdeki özniteliklerin çıkarılması amacıyla dlib kütüphanesindeki fonksiyonlar; yüz takibi, yüz tanıma ve kafa yönünün tahmini amacıyla da OpenCV kütüphanesindeki algoritmalar kullanılmıştır.

3.3.1. Yönlendirilmiş Gradyanların Histogramı (YGH)

Bu yöntem Dalal ve Triggs (2005) tarafından önerilmiş ve literatürde Histogram of Oriented Gradients (HOG) yöntemi olarak bilinmektedir. Bu yöntemin detaylarına geçmeden önce öznitelik tanımlayıcısı kavramından bahsetmek yerinde olacaktır.

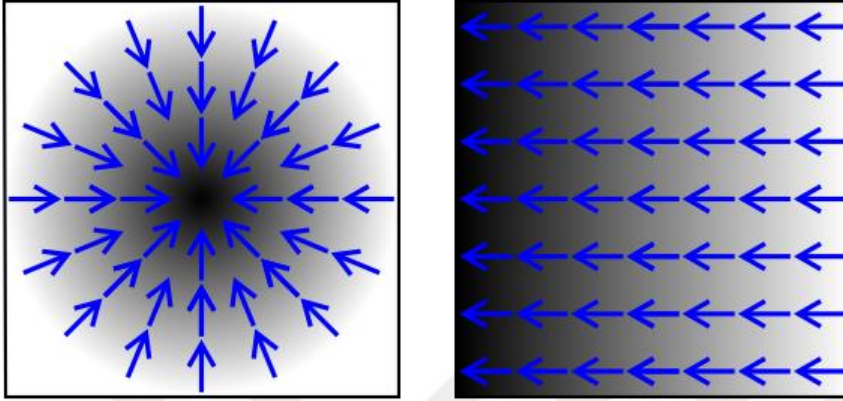
Bir görüntüyü ele alacak olursak, öznitelik tanımlayıcısı (feature descriptor) görüntüdeki gereksiz bilgilerin atılarak o görüntüyü yine aynı görüntüden çıkarılmış olan yararlı bilgiler ile ifade etmeye yarar diyebiliriz.

Renkli bir görüntü üzerinde çalışıyorsak bu görüntünün her bir pikseli 3 renk kanalının (kırmızı, yeşil, mavi) değeri ile ifade edilir. Görüntüdeki piksel sayısı ise görüntünün genişliği ile görüntünün yüksekliğinin çarpımı ile bulunur. Dolayısıyla görüntüdeki toplam bilgi $görüntü_genişliği \times görüntü_yükseklığı \times 3$ olarak ifade edilebilir. Öznitelik tanımlayıcısı bu toplam bilgi içerisindeki gereksiz kısımları atarak görüntüyü n elemanlı bir öznitelik vektörüne dönüştürmeye yarar. Öznitelik vektörü ise görüntünün görsel anlamda temsil edilmesi için değil makine öğrenme algoritmaları ile görüntünün sınıflandırılmasında, objelerin tanınmasında kullanılır (Mallick, 2016).

Öznitelik çıkarma aşamasında önemli olan kısım görüntüdeki yararlı bilginin korunmasıdır. Örneğin, kazak, tişört ve benzeri görüntü kümesi içerisindeki düğmeleri tespit etmeye çalıştığımızı düşünelim. Bir düğme genellikle daire şeklindedir ve üzerinde dikiş dikiş dikilemek için delikler bulunur. Bu durumda kenar bulma algoritması ile görüntüyü işleyip oluşan görüntüden düğmeyi bulabiliriz. Dolayısıyla kenar bilgisi yararlı bir bilgidir diyebiliriz, renk bilgisi ise gereksiz bir bilgi olduğu için atılabilir. Ayrıca öznitelikler ayırt edici olmak durumundadır. Kenar bulma algoritması düğme için de bozuk para için de benzer sonuç verir. Kullanacağımız öznitelikler bu farkı da ayırt edebilecek şekilde güçlendirilmelidir (Mallick, 2016).

Dalal ve Triggs tarafından önerilen Yönlendirilmiş Gradyanların Histogramı öznitelik tanımlayıcısında görüntüdeki yoğunluk değişimlerinin yönlerine göre oluşturulan histogramlar öznitelik olarak kullanılmaktadır. Önermiş oldukları yöntemde $64 \times 128 \times 3$ 'lük bir görüntü 3780 elemanlı bir öznitelik vektörü ile temsil edilmektedir (Mallick, 2016).

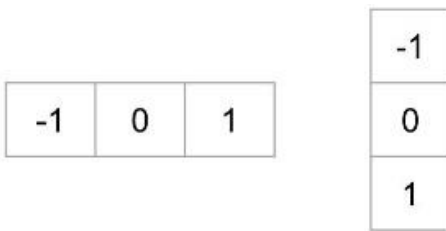
Bir görüntü işleme kavramı olan gradyan, görüntüdeki yoğunluğu ve renk değişim yönünü belirtir (Şekil 3.1).



Şekil 3.1. Gradyanın yönü (mavi oklar) ve gradyanın büyüklüğü (açık-koyu bölgeler) (Wikipedia-Image Gradient)

Gradyanların oluşturulması

Bir görüntünün YGH özniteliklerinin çıkarılması için ilk önce görüntü istenilen boyutlara getirilir ve Şekil 3.3'teki maskeler ile filtrelenerek x ve y eksenindeki gradyan görüntüleri elde edilir. Bu maskelere alternatif olarak filtrelemede Sobel operatörü de kullanılabilir.



Şekil 3.2. x ve y eksenindeki gradyan görüntülerinin elde edilmesi için kullanılan maskeler.

x ve y eksenlerinde filtrelenmiş yeni görüntülerimiz sırasıyla g_x ve g_y olmak üzere gradyanın büyüklüğü (g) ve yönü (θ) Eş. 3.1'deki gibi hesaplanır:

$$g = \sqrt{g_x^2 + g_y^2} \quad (3.1)$$

$$\theta = \arctan \frac{g_y}{g_x}$$

Bir görüntünün x ve y eksenindeki yoğunluk değişimleri (kenarları ve köşeleri) görüntünün gradyanı ile belirgin hâle gelir. Örnek bir kedi görüntüsünün x ve y eksenindeki gradyan görüntüleri Resim 3.2’de verilmiştir. x eksenindeki gradyanın dikey çizgileri, y eksenindeki gradyanın ise yatay çizgileri belirginleştirdiği gözlemlenebilir. Ayrıca gradyanın görüntüdeki gereksiz birçok bilgiyi atıp görüntüdeki hatları ve yoğunluk değişimlerini vurguladığı söylenebilir.



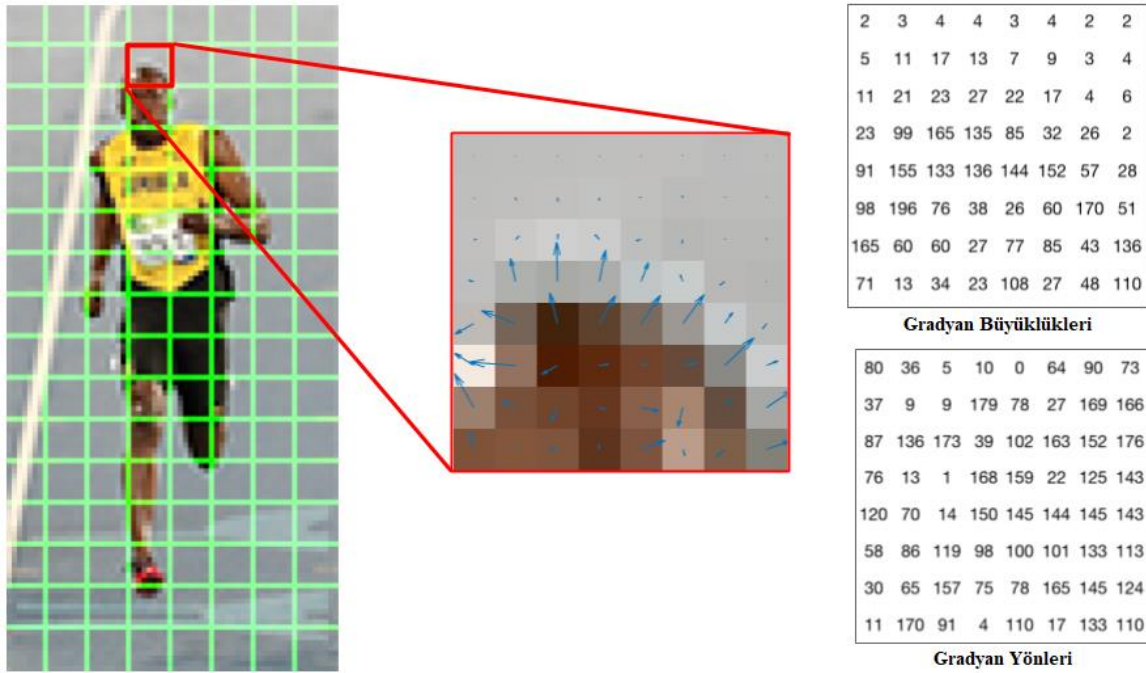
Resim 3.2. Orijinal görüntü (solda), x ekseninde gradyan görüntü (ortada), y ekseninde gradyan görüntü (sağda)

Gradyanların histogramlarının hesaplanması

Her bir piksel için gradyan büyüklükleri (g) ve yönleri (θ) elde edildikten sonra ikinci aşamada görüntü bloklara ayrılarak her bir blok için gradyanların histogramları hesaplanır.

Dalal ve Triggs bu yöntemi yayaların tespiti amacıyla kullanmış ve 64 piksel genişliğinde, 128 piksel yüksekliğinde görüntüler üzerinde çalışmıştır. Blokları ise görüntüdeki detayları ayırt edebilecek şekilde 8 piksel genişliğinde ve yüksekliğinde belirlemiştir.

Şekil 3.3’te örnek bir görüntü (solda) ve bu görüntüdeki bir görüntü bloğu verilmiştir. Bu görüntü bloğuna ilişkin gradyanların büyüklük ve yön matrisleri (sağda) verilerek gradyanların yönleri okların yönüyle, büyüklükleri ise okların uzunlukları ile ifade edilmiştir (ortada).



Şekil 3.3. Her bir 8 x 8 piksellik görüntü bloğu için gradyanların çıkarılması¹ (Mallick, 2016)

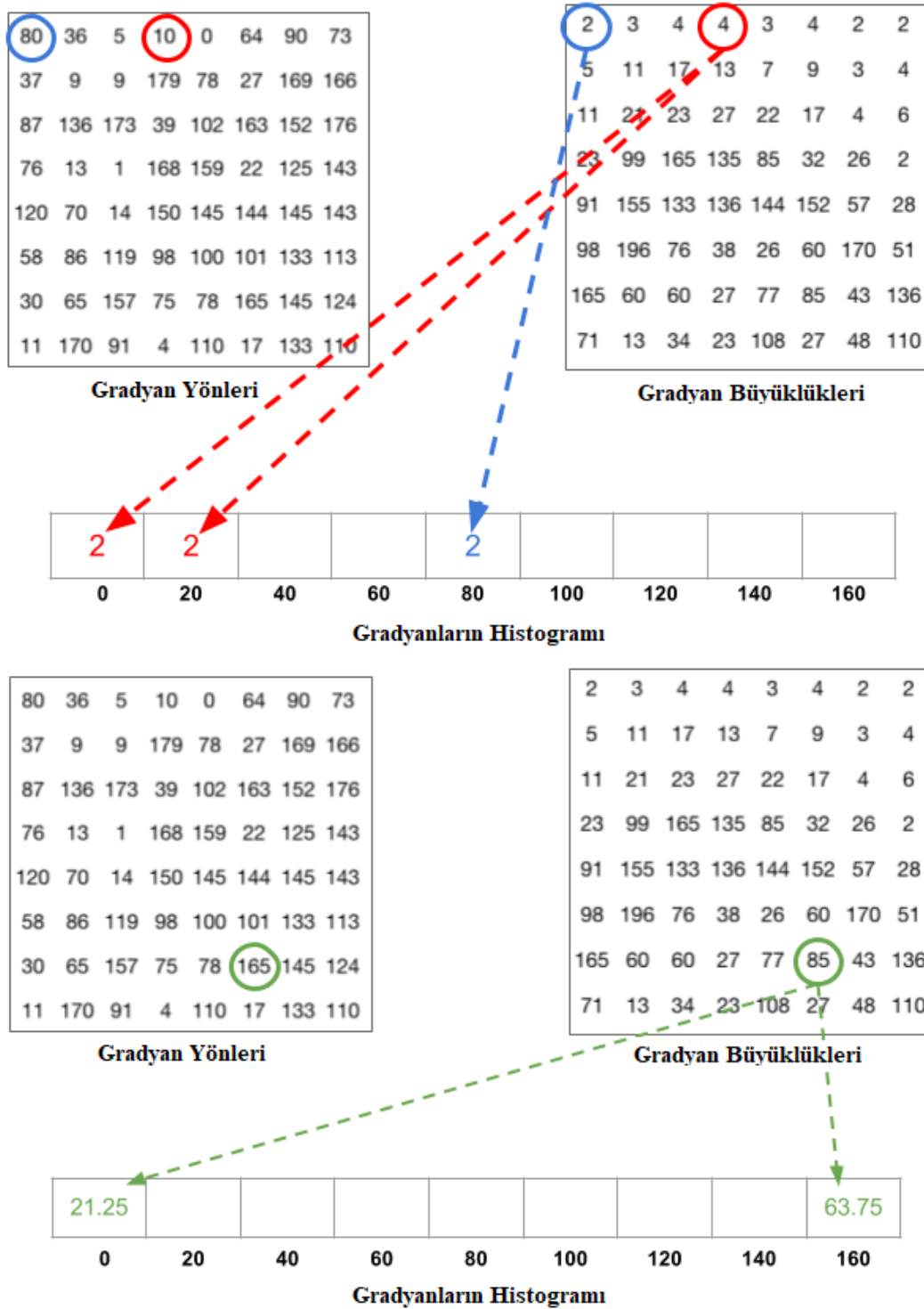
Gradyan yön matrisinde dikkat çeken husus açıların 0-360° yerine 0-180° arasında ifade edilmiş olmasıdır. Bu gösterim “işaretsiz gradyanlar” (unsigned gradients) olarak bilinmekte ve birbiriyle 180° zıt yönü gösteren gradyanların aynı açı değeri ile ifade edilmesini sağlamaktadır. Yapılan deneyler işaretsiz gradyanların yaya tespitinde daha başarılı olduğunu ortaya koymuştur.

Görüntü blokları için gradyanlar çıkarıldıktan sonra her bir blok için histogramlar oluşturulur. Histogramda 0°, 20°, 40°, 60°, 80°, 100°, 120°, 140° ve 160° olmak üzere 9 eleman bulunmaktadır.

Şekil 3.4’de bir görüntü bloğuna ait verilmiş olan gradyan yönleri ve gradyan büyüklükleri matrisinden histogramın nasıl elde edildiği gösterilmektedir. Mavi dairedeki pikselin gradyan yönünün 80° ve gradyan büyüklüğünün 2 olduğu görülmektedir. Öyleyse histogramdaki 80° olan elemana 2 değeri eklenir. Kırmızı dairedeki pikselin gradyan yönü 10° ve gradyan büyüklüğü 4 olduğundan gradyan büyüklüğü histogramdaki 0° ve 20° olan elemanlara eşit olarak ilâve edilir. Yeşil daireyle gösterilen pikselde ise durum diğerlerinden biraz farklıdır. Bu pikselin gradyan yönü 165° ve gradyan büyüklüğü 85 olarak

¹ Görsel ve görsele ilişkin açıklamalar Satya Mallick’in izniyle “Histogram of Oriented Gradients” başlıklı yazısından alınmıştır.

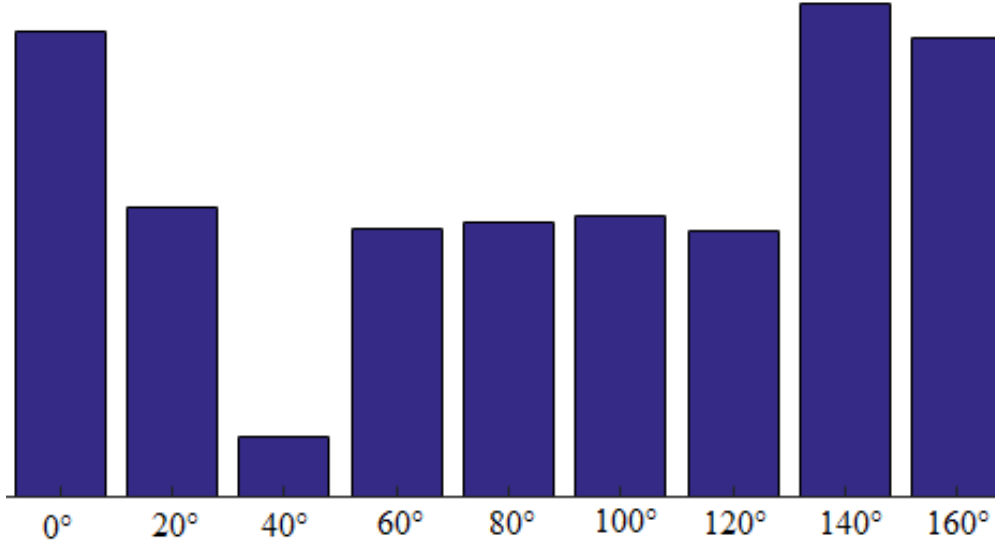
görülmektedir. İşaretsiz gradyanlarda 180° ile 0° aynı açığı ifade ettiği için 165° 'lik bir gradyana ait büyüklük değeri 160° ve 0° histogram elemanlarına, onlara olan mesafeleri ile doğru orantılı olarak paylaşılır.



Şekil 3.4. Gradyanların histogramının oluşturulması¹ (Mallick, 2016)

¹ Görsel ve görsele ilişkin açıklamalar Satya Mallick'in izniyle "Histogram of Oriented Gradients" başlıklı yazısından alınmıştır.

Yukarıdaki görüntü bloğu için verilen gradyan büyüklüklerinin ilgili açılara paylaştırılması sonucunda histogram elde edilir (Şekil 3.5).



Şekil 3.5. Gradyanların histogramı¹ (Mallick, 2016)

Gradyanların normalizasyonu

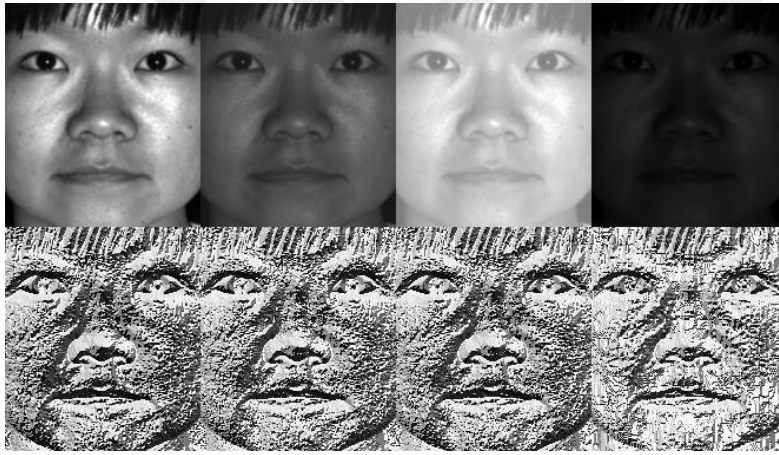
Bir görüntünün gradyanları ışıklandırma değişimlerine duyarlıdır. YGH öznitelik tanımlayıcısının bu değişimlerden etkilenmemesi için 16 x 16 piksel büyüklüğündeki ve 4 histogramı kapsayan yeni bloklar toplu olarak normalize edilir. 4 histogram 36 x 1 elemanlı bir vektörle ifade edilebilir. Normalizasyon yapılırken vektörün uzaydaki uzunluğu hesaplanarak bu uzunluk değeri vektörün tüm elemanlarına bölünür. 16 x 16 piksellik bloktaki normalizasyon tamamlandıktan sonra görüntünün genişliği boyunca her seferinde 8 piksel sağa kayıp tekrardan 16 x 16'lık blok için normalizasyon gerçekleştirilir. Görüntünün genişliğinin sonuna ulaşıldığında görüntünün en sol tarafına tekrar geçilip 8 piksel aşağı kayarak elde edilen 16 x 16'lık blok için normalizasyon yapılır. Bu işlem tüm görüntü dolaşılacak şekilde devam eder. Normalizasyon tamamlandıktan sonra her bir 16 x 16'lık bloğa ait 36 x 1'lik histogram vektörleri birleştirilerek görüntüyü temsil edecek tek bir YGH öznitelik vektörü oluşturulmuş olur. Oluşturulan bu öznitelik vektörü objelerin tanınması gibi işlerde kullanılmaktadır.

¹ Görsel ve görsele ilişkin açıklamalar Satya Mallick'in izniyle "Histogram of Oriented Gradients" başlıklı yazısından alınmıştır.

3.3.2. Yerel İkili Örüntüler (YİÖ)

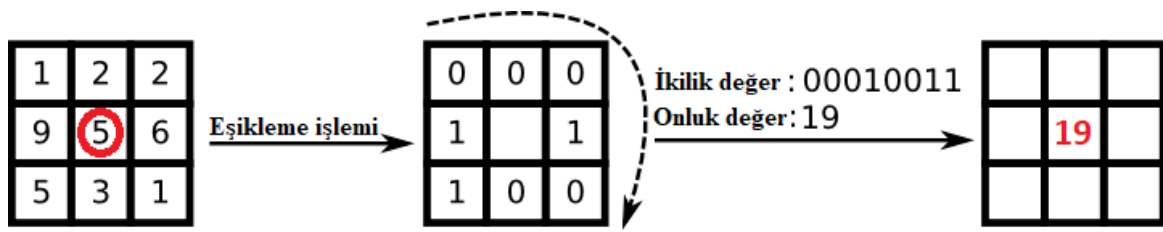
Yerel İkili Örüntüler (Local Binary Patterns) yöntemin temelleri ilk olarak Ojala ve diğerleri (1994, 1996) tarafından atılmıştır. Yerel İkili Örüntüler yöntemi, örüntü tanıma uygulamalarında yaygın olarak kullanılmakta olan piksellerin yerel komşuluğuna dayalı bir yöntemdir.

Yüz tanımda yaygın olarak kullanılan diğer yöntemler olan Özyüzler (Eigenfaces) ve Fisherfaces'in aksine YİÖ yöntemi bütüncül (holistic) bir yöntem değildir, özneliklerin istatistiksel olarak çıkarımını sağlar. Işık değişimlerinden çok etkilenmemesi de bu yöntemin güçlü yanlarından (Resim 3.3).



Resim 3.3. Yerel ikili Örüntüler yönteminin ışıklandırma değişimi karşısında başarısı (OpenCV-LBPH)

Bu yöntem temel olarak gri tonlamaya çevrilmiş bir görüntüdeki her bir pikselin komşu piksellerine göre hesaplanmış olan bir ikili değer ile ifade edilmesine dayanır. YİÖ'nün orijinal sürümünde, bir görüntüdeki 3x3'lük piksel büyüklüğündeki blokların merkezindeki piksel değeri eşik değer olarak belirlenir ve bu değer komşu pikseller ile tek tek kıyaslanır. Komşu pikselin değerinin eşik değerinden küçük olması durumunda komşu pikselin yeni değeri "0", tersi durumda ise "1" olarak atanır. 8 komşu piksel için bu işlem tamamlandıktan sonra belirlenen yöne (saat yönü ya da saat yönünün tersine olacak şekilde) göre komşu pikseller sırasıyla alınarak 8 basamaklı bir ikili sayı elde edilir. İkilik sistemdeki bu sayının onluk karşılığı merkez pikselin YİÖ kodu (LBP code) olmuş olur. Görüntüdeki tüm pikseller bu şekilde işlenerek LBP görüntüsü elde edilir. Örnek bir eşikleme işlemi Şekil 3.6'da görülmektedir.



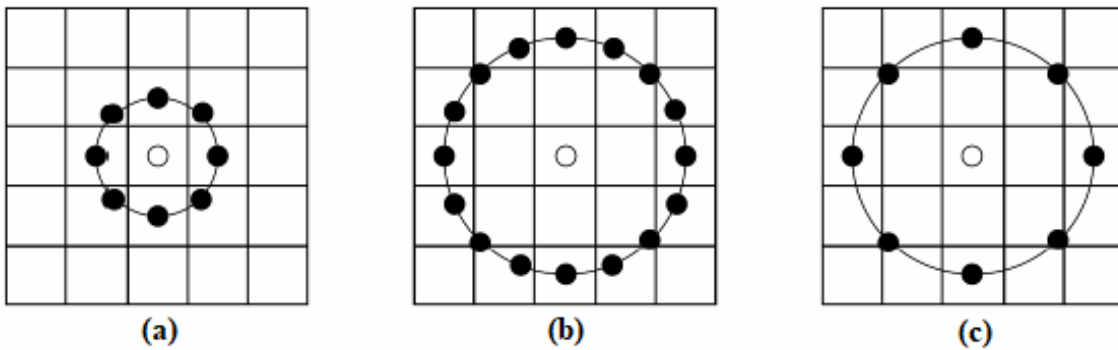
Şekil 3.6. Yerel ikili Örüntüler yönteminin uygulanması

(x_m, y_m) 2 boyutlu görüntüde merkezdeki pikselin koordinatı, d_m merkezdeki pikselin değeri ve d_p işlem yapılan komşu pikselin değeri olmak üzere yukarıda anlatılan işlem matematiksel olarak Eş. 3.2'deki gibi ifade edilebilir:

$$YİÖ(x_m, y_m) = \sum_{p=0}^{p-1} 2^p \times f(d_p - d_m) \quad (3.2)$$

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Sabit 3x3'lük bloklar ile işlem yapılan orijinal YİÖ yönteminin dokuların (texture) sınıflandırmasında başarılı olduğu, fakat farklı ölçeklerde bu yöntemin başarılı sonuçlar vermediği gözlemlenmiştir. Ahonen ve diğerleri (2006) yapmış oldukları çalışma ile bu yöntemi yüz tanıma uygulamalarında kullanarak geliştirmiş ve yöntemin başarısını ortaya koymuştur. YİÖ'nün geliştirilmiş sürümü olarak kabul edilen ve Dairesel YİÖ (Circular LBP) olarak da anılan yeni yöntem, ölçeklendirme problemini aşmak için yarıçap (r) ve piksel sayısı (n) olmak üzere iki yeni parametre ile mevcut yöntemi daha kararlı bir şekilde yeniden ifade etmiştir. Dairesel YİÖ'daki komşuluk örnekleri Şekil 3.7'de görülmektedir.



Şekil 3.7. Dairesel YİÖ örnekleri. (a) $r = 1, n = 8$ (b) $r = 2, n = 16$ (c) $r = 2, n = 8$

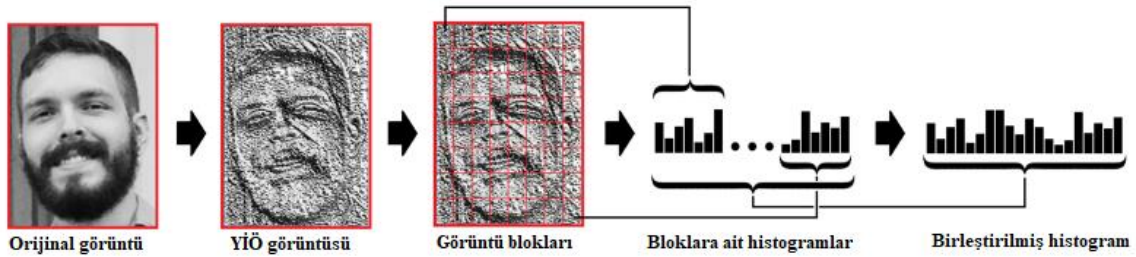
YİÖ yarıçapı r , YİÖ komşu piksel sayısı n , işlem yapılacak olan komşu piksel p , merkezdeki pikselin koordinatı (x_m, y_m) , işlem yapılacak olan komşu pikselin koordinatı (x_p, y_p) ve $p \in n$ olmak üzere Dairesel YİÖ için komşu pikselin koordinatı Eş. 3.3'deki gibi bulunabilir:

$$\begin{aligned} x_p &= x_m + r \cos \frac{2\pi p}{n} \\ y_p &= y_m + r \sin \frac{2\pi p}{n} \end{aligned} \quad (3.3)$$

Bu şekilde bulunan koordinatın 2 boyutlu görüntü üzerinde belli bir piksele tekabül etmemesi durumunda Eş. 3.4'deki gibi çift yönlü doğrusal enterpolasyon (bilinear interpolation) uygulanarak koordinat bulunabilir:

$$f(x, y) = [1 - x \ x] \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1 - y \\ y \end{bmatrix} \quad (3.4)$$

Her bir bloktaki pikseller için onluk YİÖ kodları bulunup yeni YİÖ görüntüsü oluşturulduktan sonra yeni görüntüdeki her bir blok için YİÖ histogramları oluşturulur. İşlem yapılan blok içerisindeki her bir pikselin değeri ilgili histogram elemanına eklenmek suretiyle o blok için histogram elde edilir. Bu işlem görüntüdeki tüm bloklar için tekrarlanır. Son olarak elde edilen tüm histogramlar birleştirilerek görüntüyü temsil edecek olan nihai histogram vektörü elde edilmiş olur. YİÖ uygulanmış örnek bir görüntü ve bu görüntüye ait birleştirilmiş histogram Şekil 3.8'de görülmektedir.



Şekil 3.8. YİÖ histogramının oluşturulması¹ (Prado, 2017)

¹ Görsel Kelvin Salton do Prado'nun izniyle "Face Recognition: Understanding LBPH Algorithm" başlıklı yazısından alınmıştır.

3.3.3. Destek Vektör Makineleri (DVM)

Bir makine öğrenmesi yöntemi olan Destek Vektör Makineleri (Support Vector Machines - SVM) etiketli bir eğitim verisi kümesi üzerinde sınıflandırma ve regresyon yapabilen gözetimli (supervised) bir öğrenme yöntemidir.

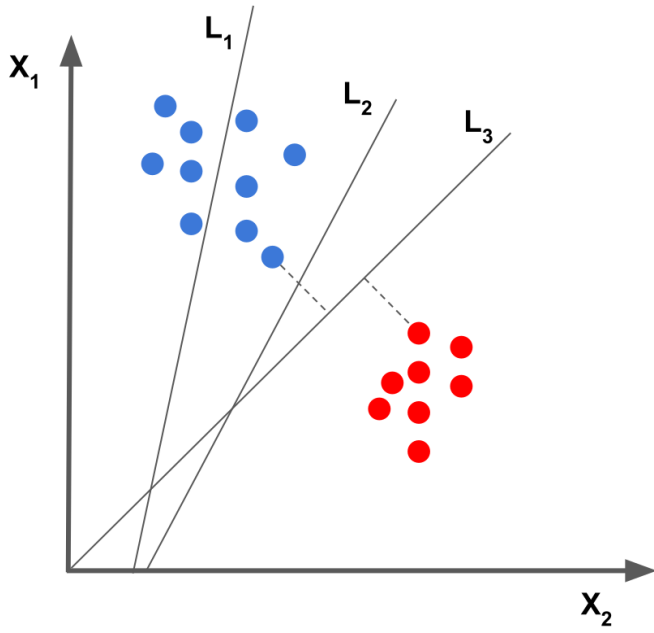
Destek Vektör Makineleri konusuyla ilgili fikirlerin birçoğu ilk olarak Rusya Bilimler Akademisi'nde Vapnik ve Chervonenkis tarafından bilgisayar öğrenimi ve örüntü tanıma maksadıyla öne sürülen “Genelleştirilmiş Portre Yöntemi” (Generalised Portrait Method) kapsamında ortaya atılmıştır. Bu fikirlerin gelişimi 1962 yılında başlamış ve ilk olarak 1964 yılında yayınlanmıştır (Chervonenkis, 2013).

Destek Vektör Makineleri'nin ivme kazanmasında Bernhard E. Boser, Isabelle M. Guyon ve Vladimir N. Vapnik'in COLT 1992 konferansındaki önerileri etkili olmuştur. Bu konferansta hâlihazırda kullanılan doğrusal DVM'lere ek olarak, maksimum marjinli hiper-düzlemlere "çekirdek hilesi" (kernel trick) uygulayarak doğrusal olmayan sınıflandırıcıların yaratılabilmesi için bir yol önermişlerdir (Boser, Guyon ve Vapnik, 1992).

1995 yılında yumuşak marjinli sınıflandırıcı (soft margin classifier) Cortes ve Vapnik (1995) tarafından öne sürülmüştür. Yine aynı yıl DVM algoritması Vapnik tarafından regresyon işlemlerinde de kullanılabilecek şekilde geliştirilmiştir (Vapnik, 1995).

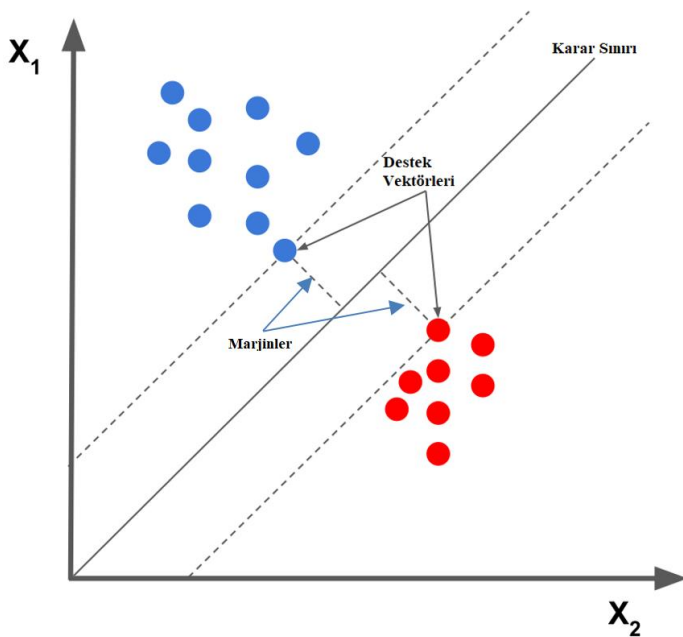
Destek Vektör Makineleri'ni anlamak adına iki farklı sınıftan oluşan, 2 boyutlu bir düzlemde ifade edilebilen ve doğrusal olarak ayrıştırılabilen bir veri setini ele alacak olursak DVM bu iki sınıfı birbirinden ayırabilecek en iyi çizgiyi bulmayı hedefler.

Şekil 3.9'da mavi noktalarla ve kırmızı noktalarla belirtilen iki ayrı sınıf ve bu sınıfları ayırmak için çizilen L_1 , L_2 ve L_3 çizgileri görülmektedir. L_1 çizgisi sınıfları birbirinden ayıramadığı için iyi bir seçim değildir. L_2 ve L_3 'ün ise sınıfları birbirinden ayırdığı, fakat L_3 'ün sınıfları daha belirgin bir şekilde ayırdığı gözlemlenmektedir.



Şekil 3.9. DVM'de sınıfları ayırmak için çizilen çizgiler (Mallick, 2018)

DVM'nin eğitilmesi aslında bu en iyi çizginin, yukarıdaki örneğe göre L_3 'ün bulunması anlamına gelir. DVM'lerde bu çizgi *karar sınırı (decision boundary)* olarak isimlendirilir. Sınıfları birbirinden ayıran ve karar sınırına en yakın olan elemanlar ise *destek vektörleri (support vectors)* olarak anılır. DVM karar sınırını belirlerken marjini maksimize eden yani destek vektörlerine mümkün olduğunca uzakta ve sınıfları birbirinden ayıracak şekilde bir iyileştirme yapar. Burada anlatılan kavramlar Şekil 3.10'da daha anlaşılır olarak görülebilir.



Şekil 3.10. DVM'de karar sınırı ve destek vektörleri

Eđitim tamamlandıktan sonra sisteme verilecek olan yeni bir nokta karar sınırının hangi tarafında olduđuna gore DVM modeli tarafından sınıflandırılır.

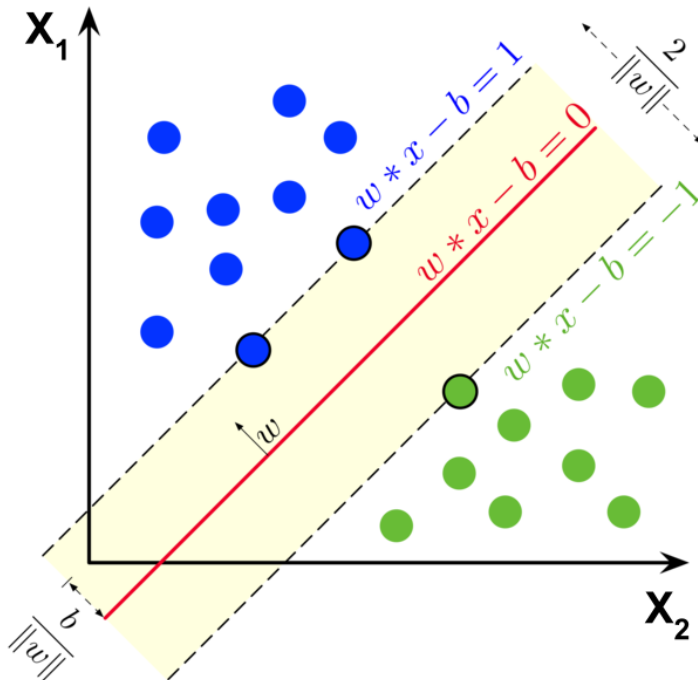
Dođrusal olarak ayrılabilen n elemanlı eđitim kumemiz matematiksel olarak $(\vec{x}_1, y_1), \dots, (\vec{x}_n, y_n)$ Őeklinde ifade edilebilir. Buradaki \vec{x}_i eđitim veri kumemizdeki her bir elemanı, y_i ise -1 ya da 1 olmak uzere bu elemanın hangi sınıfa ait olduđunu belirtmektedir. \vec{w} ađırlık vektorune (karar sınırına dik olan normal vektoru), b eđilim katsayısına karŐılıklı gelmek uzere, bir dođrusal DVM iin karar sınırı EŐ. 3.5'deki gibi ifade edilebilir:

$$\vec{w} * \vec{x} - b = 0 \quad (3.5)$$

Karar sınırını bulmakta kullanılan ve sınıfları birbirinden ayıran iki paralel dođru (negatif ve pozitif hiperduzlemler) ise EŐ. 3.6'daki gibidir:

$$\begin{aligned} \vec{w} * \vec{x} - b &= 1 \\ \vec{w} * \vec{x} - b &= -1 \end{aligned} \quad (3.6)$$

Bu anlatılanlar Őekil 3.11'de daha anlaşılır olarak gorulebilir.

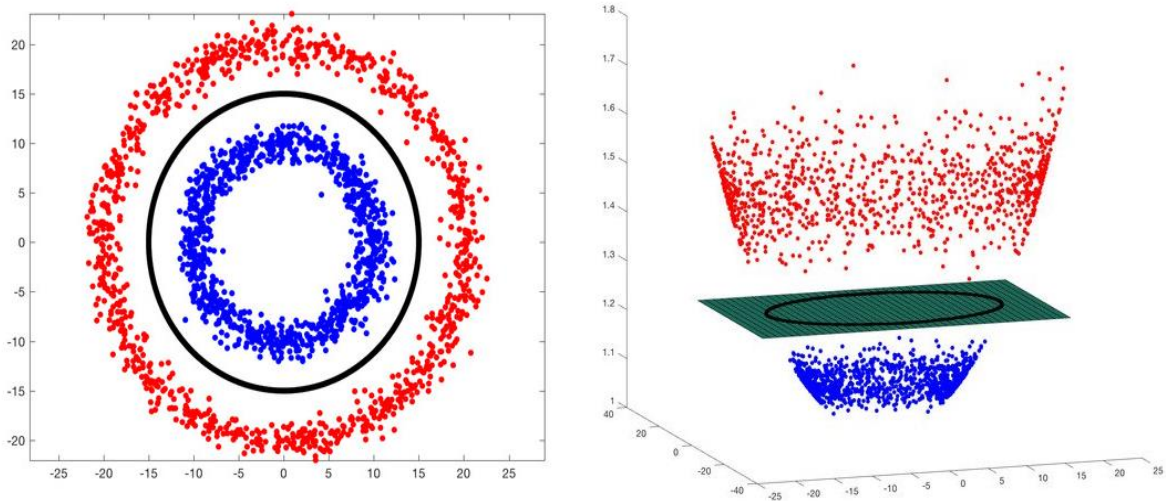


Őekil 3.11. DVM'de karar sınırının matematiksel ifadesi

DVM, örneklendirilen 2 boyutlu sınıflandırmanın yanı sıra daha fazla boyutlar için de sınıflandırma yapabilmektedir. 2 boyutlu bir veri kümesine sahipsek karar sınırımız bir çizgiyle, 3 boyutlu veri kümesine sahipsek karar sınırımız bir düzlemlerle, 3'ten fazla boyutlu bir veri kümesine sahipsek de karar sınırımız bir *hiperdüzlem* ile ifade edilir.

Ayrıca DVM yukarıdaki örneklerdeki gibi doğrusal sınıflandırmanın yanında doğrusal sınıflandırmanın mümkün olmadığı durumlarda da sınıflandırma yapabilecek şekilde geliştirilmiştir. Bu durumda “Çekirdek Hilesi” (Kernel Trick) yöntemi kullanılarak veriler daha yüksek boyutlu uzaya taşınır ve sınıflandırma yapılır. DVM ile bu amaçla birçok çekirdek fonksiyonu kullanılabilir. Bunlardan yaygın olarak kullanılanları Doğrusal (Linear), Polinom (Polynomial), Gauss Radyal Tabanlı Fonksiyon (Gaussian Radial Basis Function) ve Sigmoid çekirdek fonksiyonlarıdır.

Şekil 3.12’te doğrusal olarak sınıflandırılmayan bir veri kümesine Gauss Radyal Tabanlı Fonksiyon ile çekirdek hilesi uygulandıktan sonra doğrusal olarak sınıflandırabildiği görülmektedir.



Şekil 3.12. DVM’nin doğrusal olarak sınıflandıramadığı veri kümelerinde çekirdek hilesinin uygulanması¹ (Mallick, 2018)

¹ Görsel Satya Mallick’in izniyle “Support Vector Machines (SVM)” başlıklı yazısından alınmıştır.

3.3.4. PnP probleminin çözülmesi ve Euler açılarının bulunması

Perspective-n-Point ya da kısaca PnP problemi, 3 boyutlu uzayda n tane noktasının koordinatları bilinen bir nesnenin yine aynı noktalara ait 2 boyutlu düzlemdeki koordinatları da biliniyorsa kalibre edilmiş kameranın pozunu (pose) tahmin etme problemidir (Horaud, Conio, Le Boulleux ve Lacolle, 1989).

Bu problemin çözülmesi ile kamera karşısındaki nesnelerin kameraya göre dönme (rotation) ve öteleme (translation) değerleri tahmin edilebilir. 3 boyutlu bir nesnenin kameraya göre dönme hareketini x , y ve z eksenleri etrafındaki dönme açıları ile ifade edebiliriz. Öteleme için de aynı durum geçerlidir. Bir 3 boyutlu nesnenin kameraya göre ötelemesini bilmek için x , y ve z eksenlerine göre öteleme değerlerini bilmek gerekir. Bu durumda, dönme ve öteleme değerlerinin birleşimi olan ve literatürde poz (pose) olarak anılan bir nesnenin kameraya göre hareketi 6 değer bilmesini gerektirmektedir.

Görüntü işleme alanında araştırmacıların ilgisini çeken bu problemin çözümü adına birçok yöntem önerilmiştir. Bu çözümlerden P3P olarak anılan Gao, Hou, Tang ve Cheng (2003)'in çözümü, EPnP olarak bilinen Lepetit, Moreno-Noguer ve Fua (2009)'nın çözümü ile Schweighofer ve Pinz (2006)'in önerdiği PnP çözümleri iteratif olmayan çözümler olarak öne çıkmaktadır. Lu, Hager ve Mjolsness (2000)'in önerdiği çözüm ise iteratif çözümlerdendir.

İteratif çözümlerin daha yüksek doğrulukta sonuç verdiği, fakat düzgün bir şekilde ilklendirilmediğinde (initialize) kararsız hâle geldiği ve yakınsama işleminin uzun zaman aldığı belirtilmektedir (Lock, 2016).

OpenCV kütüphanesi içerisinde gelen solvePnP() fonksiyonu bu problemin birçok çözümünü araştırmacılara hazır olarak sunmaktadır. solvePnP() fonksiyonunu kullanabilmek için:

- objeye ait noktaların 3 boyutlu koordinatlarını,
- objeye ait görüntüdeki aynı noktalar için olan 2 boyutlu koordinatların ve
- kullanılan kameraya özgü olan parametrelerin

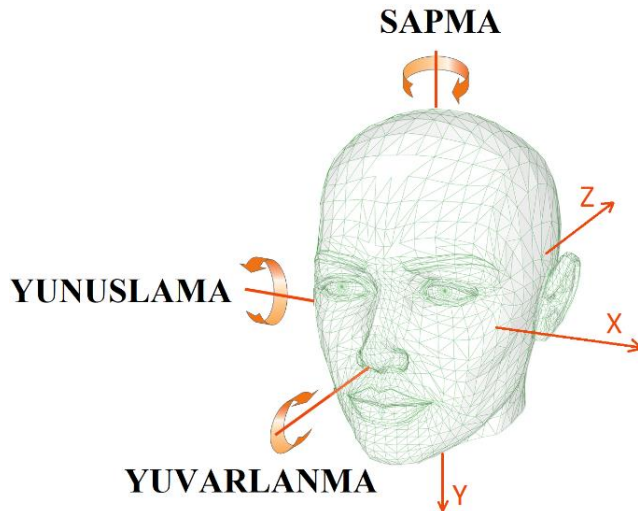
bilinmesi gerekmektedir.

Ayrıca $\text{solvePnP}()$ fonksiyonuna hangi PnP çözümünün tercih edildiği de giriş parametresi olarak verilebilmektedir. OpenCV 3 ve sonrasındaki sürümlerde bu parametreye verilebilecek olan değerler şunlardır:

- **SOLVEPNP_ITERATIVE:** $\text{solvePnP}()$ fonksiyonuna hiçbir parametre belirtilmediğinde varsayılan olarak seçilen ve Levenberg-Marquardt optimizasyonuna (Levenberg, 1944; Marquardt, 1963) dayanan OpenCV kütüphanesinin kendi PnP çözümüdür.
- **SOLVEPNP_P3P:** Gao ve diğerlerinin (2003) P3P olarak bilinen çözümüdür.
- **SOLVEPNP_EPNP:** Lepetit ve diğerlerinin (2009) EPnP olarak bilinen çözümüdür.
- **SOLVEPNP_DLS:** Hesch ve Roumeliotis (2011)'in PnP çözümüdür.
- **SOLVEPNP_UPNP:** Penate-Sanchez, Andrade-Cetto ve Moreno-Noguer (2013)'ün PnP çözümüdür.

$\text{solvePnP}()$ fonksiyonu çalıştırıldıktan sonra çıkış parametresi olarak dönme vektörünü (rotation vector) ve öteleme vektörünü (translation vector) üretir.

Euler açılarının nasıl hesaplanacağına geçmeden önce bu açıları tarif etmek yerinde olacaktır. Euler açıları ismini Leonhard Euler'den almaktadır. Leonhard Euler tarafından 1776 yılında takdim edilmiş olan bu açıları, bir katı cismin sabit bir koordinat sistemine göre yönünü tanımlayan üç açıdır (Euler, 1776). Bu üç açı yunuslama (pitch), sapma (yaw) ve yuvarlanma (roll) açıları olarak da bilinmektedir. Bir kafa örneği için Euler açıları Şekil 3.13'de görülmektedir.



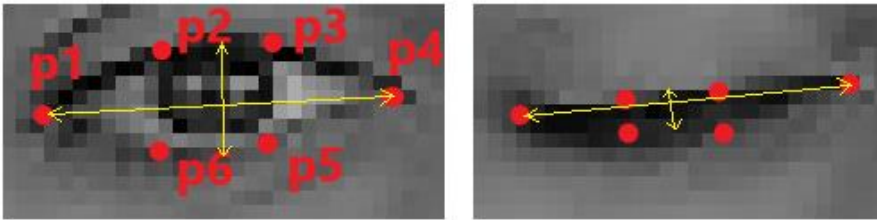
Şekil 3.13. Euler açıları

Euler açılarının nasıl hesaplanacağına gelecek olursak, öncelikle solvePnP() fonksiyonundan elde etmiş olduğumuz dönme vektörüne Rodrigues dönme formülü (Wikipedia-RRF, 2019) uygulanarak dönme matrisi elde edilir. Bu dönüşüm için OpenCV kütüphanesi içerisinde hazır olarak gelen Rodrigues() fonksiyonu kullanılabilir. Elde edilen dönme matrisi kullanılarak Slabaugh (1999)'da tarif edildiği gibi bir dönüşüm uygulanarak dönme matrisine karşılık gelen Euler açıları hesaplanır.

3.3.5. Göz kısıklığının tespit edilmesi

Göz kısıklığı, literatürde göz boy-en oranı (Eye Aspect Ratio) olarak bilinen oran ile hesaplanmaktadır. Soukupová ve Cech (2016) göz boy-en oranını göz kırpmaları tespit etmek amacıyla başarıyla kullanmıştır.

Göz kısıklığı öğrencinin uykulu ya da dikkatsiz olduğunu tespit etmek adına önemli bir parametredir. Gözler kısıldığında ya da kapandığında göz boy-en oranı 0 değerine yaklaşır. Bu durum Şekil 3.14'de görülmektedir.



Şekil 3.14. Gözün açık ve kapalı olma durumuna göre nirengi noktalarının durumu

Şekil 3.14.'de görülen nirengi noktalarının koordinatları ve Öklid uzaklığı formülü ile göze ait boy ve en bilgileri basitçe hesaplanıp boy-en oranı Eş. 3.7'deki gibi bulunabilir:

$$BEO = \frac{\|p2 - p6\| + \|p3 - p5\|}{2 \times \|p1 - p4\|} \quad (3.7)$$

Bu çalışmada dikkat dağınıklığı temelde kafa yönü açılarına dayalı olarak tespit edilmeye çalışılmış, göz kısıklığı tespiti ise öğrencilerin göz bölgesinin yeterince belirgin olduğu durumlarda dikkat dağınıklığının tespit başarısını artırmak amacıyla opsiyonel olarak eklenmiştir.

3.4. Kullanılan Veri Kümesi

Geliştirilen uygulamada kafa yönü tahmini, yüz tanıma ve dikkat dağınıklığı tespitinin başarısının test edilebilmesi amacıyla UPNA Head Pose Database (Ariz, Bengoechea, Villanueva ve Cabeza, 2016) kullanılmıştır.

3.4.1. UPNA Head Pose Database

Bu veritabanı 6'sı erkek 4'ü kadın 10 ayrı kişiye ait videolardan oluşmaktadır. Kafa takibi ve kafa yönü ile ilgili çalışmalarda kullanılması amacıyla tasarlanmıştır. Veritabanı içerisinde her bir kişi için çekilen 12 adet video dosyası olmak üzere toplamda 120 adet video dosyası bulunmaktadır. Bu videolardaki kişiler kafalarını x, y, z eksenleri boyunca hareket ettirmiş (translation) ve döndürmüşlerdir (rotation). Veritabanındaki videolar 1280 x 720 çözünürlüğünde 30fps ile MPEG-4 biçiminde kaydedilmiş olup videolardan her biri 10 saniyeden, yani 300 kareden (frame) oluşmaktadır (Ariz ve diğerleri, 2016).

Veritabanıyla birlikte her bir video için sensörler vasıtasıyla elde edilen ölçüm değerlerini içeren metin belgeleri de paylaşılmıştır. Her bir video için 3 tip metin belgesi paylaşılmıştır:

1. **2 boyutlu ölçüm değerleri (*_groundtruth2D.txt):** Her bir kare için bir satır olmak üzere 300 satırdan oluşur. Her bir satırında TAB karakteri ile ayrılmış 108 sütun bulunur. Bu 108 kolonda yüzdeki 54 nirengi noktasına ait x ve y koordinat değerleri (x1 y1 x2 y2 ... x54 y54 şeklinde) bulunur.
2. **3 boyutlu ölçüm değerleri (*_groundtruth3D.txt):** Her bir kare için bir satır olmak üzere 300 satırdan oluşur. Her bir satırında TAB karakteri ile ayrılmış 6 sütun bulunur. Bu sütunlarda kafanın x, y ve z eksenleri boyunca öteleme (translation) miktarları mm cinsinden ve dönme (rotation) açıları derece cinsinden verilmiştir. Bu değerlerin kullanıcının kafasına takılmış olan sensör yardımıyla yüksek hassasiyetle ölçüldüğü bildirilmiştir (Ariz ve diğerleri, 2016). Bu ölçüm değerleri sırasıyla şöyledir:
 - a. Kafanın x eksenini boyunca öteleme (translation) miktarı (T_x),
 - b. Kafanın y eksenini boyunca öteleme miktarı (T_y),
 - c. Kafanın z eksenini boyunca öteleme miktarı (T_z),
 - d. Kafanın z eksenini boyunca dönme açısı olan yuvarlanma (*roll*),

- e. Kafanın y eksenine boyunca dönme açısı olan sapma (*yaw*),
- f. Kafanın x eksenine boyunca dönme açısı olan yunuslama (*pitch*)

3. 3 boyutlu yüz modeli (*_model3D.txt): 54 satırdan ve her bir TAB karakteri ile ayrılmış 3 sütundan oluşur. Bu dosya ilgili kişi için yüzdeki 54 nirengi noktasının x, y ve z koordinatlarını içerir (Ariz ve diğerleri, 2016).

Veritabanında bulunan videolardan alınmış örnek görüntüler Resim 3.4’de görülmektedir.



Resim 3.4. UPNA Head Pose Database’den alınan örnek görüntüler

Bu çalışmada bu veritabanıyla birlikte gelen 3 boyutlu ölçüm değerlerinden yuvarlanma (roll), sapma (*yaw*) ve yunuslama (*pitch*) açıları, görüntü işleme teknikleri kullanarak hesapladığımız yuvarlanma, sapma ve yunuslama açılarının doğrulanması amacıyla kullanılmıştır (Bkz. Bölüm 5.1).

UPNA Head Pose Database ayrıca dikkat dağınıklığı sınıflandırıcılarının eğitilmesinde kullanılacak olan 1000 adet etiketlenmiş fotoğraf içeren veri kümesinin oluşturulmasında da kullanılmıştır (Bkz. Bölüm 5.3).

4. DİKKAT DAĞINIKLIĞI TESPİT SİSTEMİ (DDTS) YAZILIMI

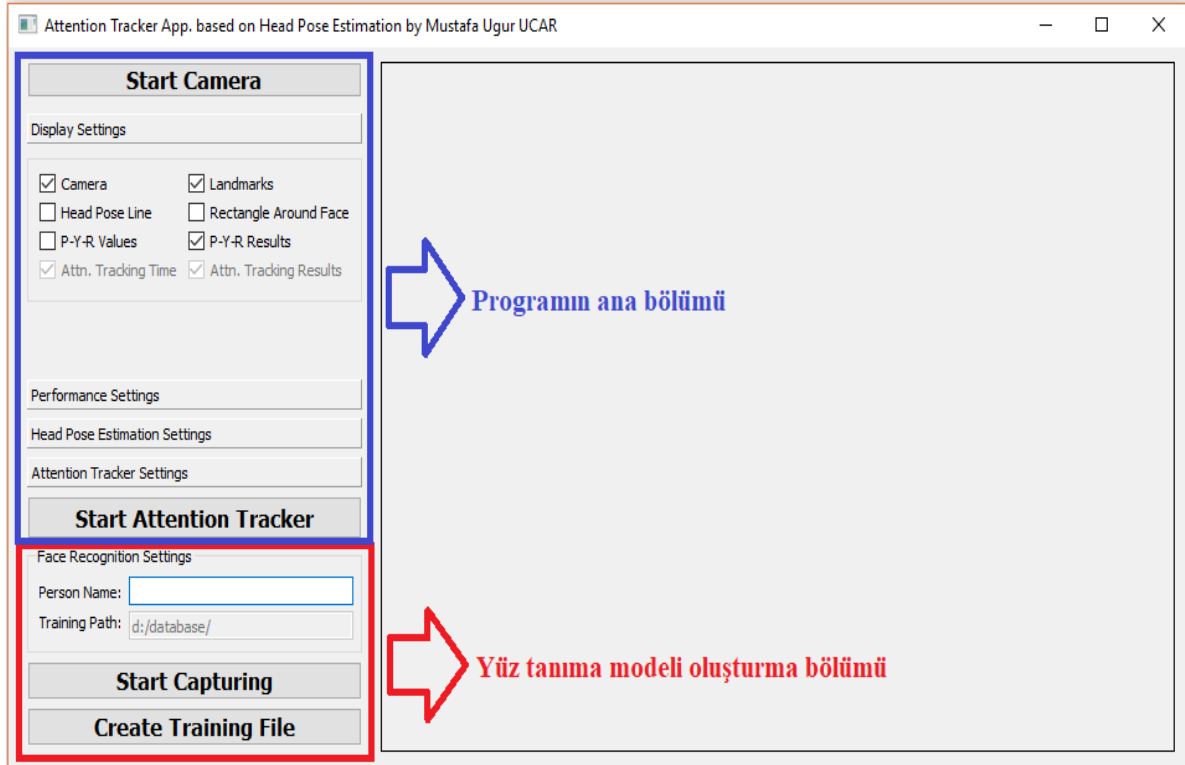
Kafa yönlerine göre dikkat dağınıklığını gerçek zamanlı olarak tespit etmek üzere tasarlanmış olan Dikkat Dağınıklığı Tespit Sistemi:

- Gerçek zamanlı olarak birden fazla öğrenciye ait yüzleri bulmakta,
- Daha önce yüz resimleri kaydedilmiş olan öğrencileri tanımakta,
- Öğrencilerin kafa yönlerini tahmin etmekte,
- Gözleri kısık olan öğrencileri tespit etmekte,
- Öğrencilerin kafa yönlerine ve göz kısıklığına göre dikkat dağınıklıklarını tespit etmektedir.

DDTS yazılımı iki bölümden oluşmaktadır. Bu bölümler şunlardır:

1. Yüz tanıma modeli oluşturma bölümü
2. Ana bölüm

DDTS yazılımına ait ekran görüntüsü ve bu bölümlerle ilgili ayarlamaların yapıldığı kısım Resim 4.1’de görülmektedir.



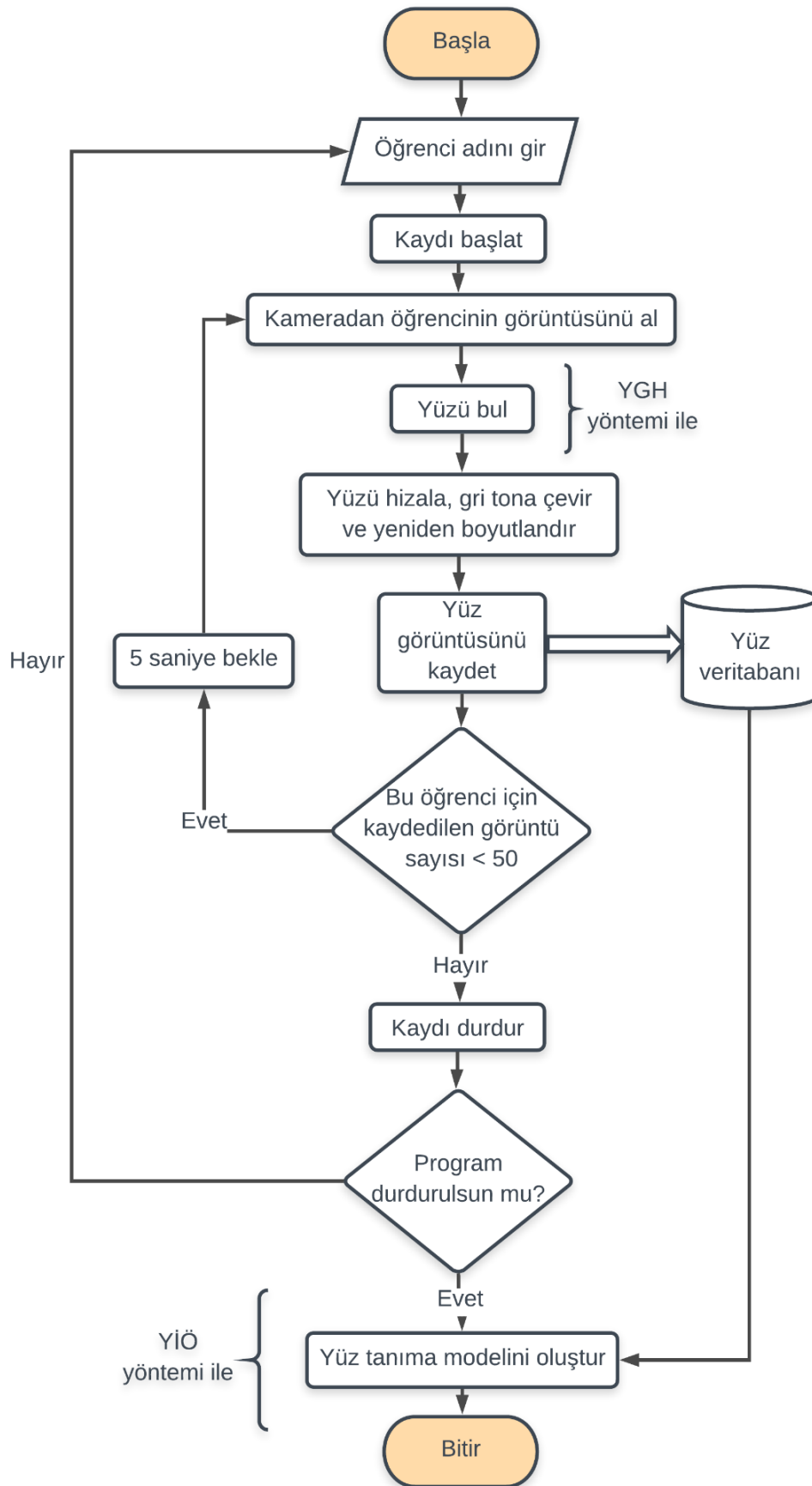
Resim 4.1. DDTS yazılımının ekran görüntüsü

4.1. Yüz Tanıma Modeli Oluşturma Bölümü

DDTS yazılımının yüz tanıma modeli oluşturma bölümünün amacı, öğrencilerin tanınabilmesini sağlayacak olan yüz tanıma modelini kolay bir şekilde oluşturmaktadır. Yüz tanıma modelini oluşturabilmek için öncelikle ismi girilen öğrencinin 5 saniye aralıklarla fotoğrafları çekilmekte ve bu fotoğraf öğrenci ismiyle oluşturulan klasöre kaydedilmektedir. Tüm öğrenciler için kayıt işlemi tamamlandıktan sonra ise bir yüz tanıma model dosyası oluşturulabilmektedir.

DDTS'nin yüz tanıma modeli oluşturma bölümünün akış diyagramı Şekil 4.1'de görülmektedir.

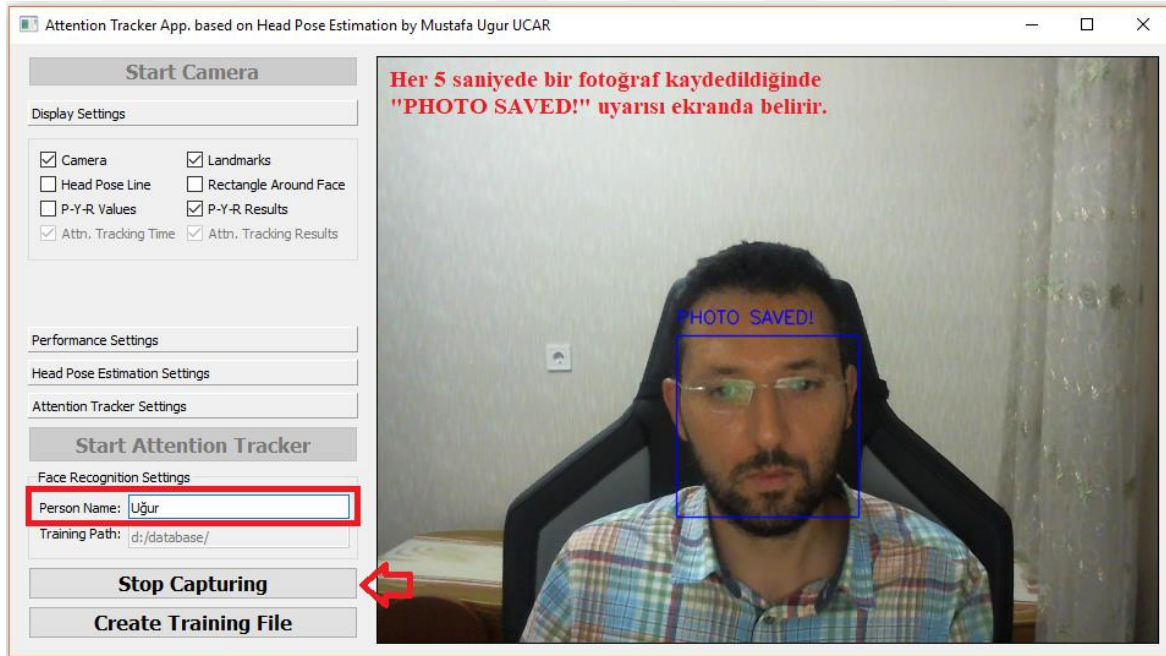




Şekil 4.1. DDTS'nin yüz tanıma modeli oluşturma bölümünün akış diyagramı

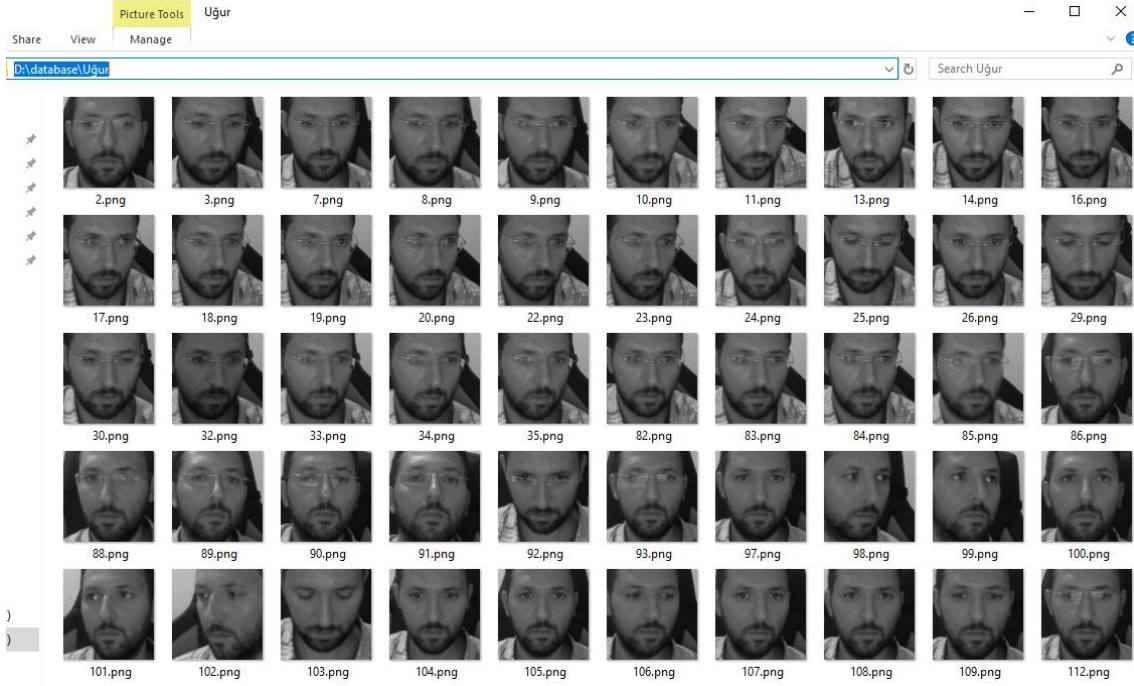
Akış diyagramında da görüleceği üzere kayıt işlemi başlatılmadan önce bir öğrenci ismi girilmesi gerekmekte ve sonrasında “Start Capturing” butonuna basılarak kayıt işlemi başlatılmaktadır (butona basıldığında buton otomatik olarak “Stop Capturing” moduna geçmektedir). Bir öğrenci için 50 adet fotoğraf çekilmekte ve kayıt işlemi sırasında öğrencinin mümkün olduğunca değişik poz vermesi (farklı yönlere bakma, kafayı döndürme, gözlüklü, gözlüksüz, gözü kapalı, gözü açık, farklı mimikler gibi) istenmektedir. Bu sayede, oluşturulacak olan yüz tanıma modelinin daha güçlü hâle gelmesi ve yüz tanıma başarısının artırılması hedeflenmektedir.

Kayıt işlemi esnasında alınmış bir ekran görüntüsü Resim 4.2’de görülmektedir.



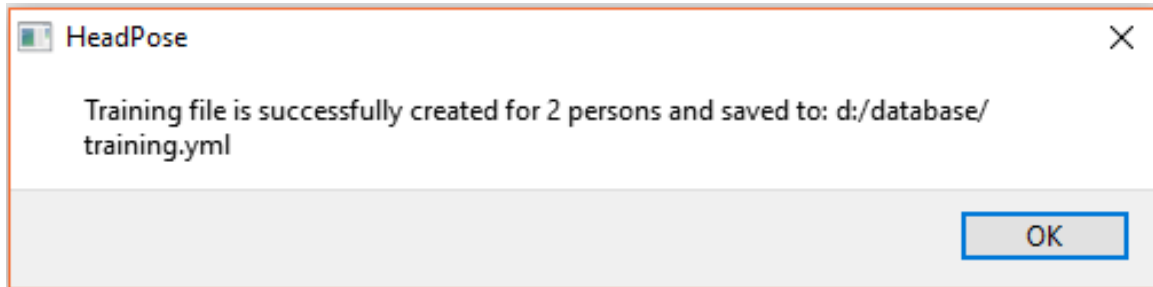
Resim 4.2. DDTS yazılımı yüz tanıma modeli oluşturma ekran görüntüsü – Yüz fotoğrafı kayıt işlemi

Kayıt işlemi sonrasında oluşan klasör ve kaydedilmiş olan fotoğraflar Resim 4.3’de görülmektedir.



Resim 4.3. DDTS yazılımı yüz tanıma modeli oluşturma – Kaydedilen örnek fotoğraflar

Tüm öğrenciler için kayıt işlemi tamamlandıktan sonra “Create Training File” butonuna basılarak yüz tanıma model dosyası oluşturulur. Bu işlem tamamlandıktan sonra ekranda Resim 4.4’deki gibi bir uyarı çıkar.



Resim 4.4. DDTS yazılımı yüz tanıma modeli oluşturma

Programın ana bölümünde burada oluşturmuş olduğumuz yüz tanıma modeli dosyası (training.yml) kullanılır ve gerçek zamanlı olarak yüz tanıma işlemi gerçekleştirilir.

4.2. Programın Ana Bölümü

DDTS'nin ana bölümünde kameradan alınan gerçek zamanlı görüntüden yüz bulma, yüzdeki nirengi noktalarının (landmark) çıkarılması, yüz tanıma, kafa yönü ve kafa yönüne dayalı dikkat dağınıklığı tahminleri yapılmaktadır.

Bunlara ilâveten eğer öğrenci kameraya çok uzak değilse ve göz bölgesi görüntü işleme ile düzgün şekilde tespit edilebiliyorsa dikkat dağınıklığı tahmin başarısını artırmak adına göz kısıklığı tespiti de yapılmaktadır. Bu özellik programın gerekli e-öğrenme ortamlarında da kullanılabilmesi adına opsiyonel olarak eklenmiştir. Sınıf ortamında yapılacak testlerde çözünürlük yeterli olmayabileceğinden bu özellik devre dışı bırakılabilir.

DDTS'nin ana bölümü ile ilgili ayarlamalar 4 kısımda toplanmıştır:

- 1) **Görüntüleme ayarları (display settings):** Kamera görüntüsünün, kafa yönü çizgisinin, tahmin edilen yunuslama (pitch) – sapma (yaw) – yuvarlanma (roll) açı değerlerinin, yüz nirengi noktalarının (landmarks), yüz etrafına çizilen dikdörtgen ile yüz tanıma sonucunun ve yunuslama-sapma-yuvarlanma sonuçlarının filtrelenebildiği kısımdır. Yunuslama için sonuç değerleri yukarı (up), aşağı (down) ya da nötr (neutral); sapma için sonuç değerleri sağ (right), sol (left) ya da nötr (neutral); ve yuvarlanma için sonuç değerleri sağa yatık (rolled right), sola yatık (rolled left) ya da nötr (neutral) şeklindedir.
- 2) **Performans ayarları (performance settings):** Kullanılan donanıma göre programın yapılandırılabilmesini sağlayan bölümdür. Kameramız çok yüksek çözünürlükte görüntü yakalamaya imkân veriyorsa programda değişiklik yapmadan yakalanacak olan görüntüde işlem yapmadan önce çözünürlüğü düşürmek için gerekli ayarlama buradan yapılabilmektedir. Ayrıca bütün karelerin işlenmesi gerçek zamanlı olarak çalışacak programımızı yavaşlatabileceğinden işlemci gücümüz de yetersiz kalıyorsa kaç karede bir işlem yapılabileceği de ayarlanabilmektedir.
- 3) **Kafa yönü tahmini ile ilgili ayarlar (head pose estimation settings):** Görüntü işleme ile elde edilen öğrencinin gerçek zamanlı yunuslama, sapma ve yuvarlanma açı değerlerine göre ekranda anlık olarak bu açı değerleri için sonuçlar gösterilmektedir (yunuslama için yukarı, aşağı ya da nötr gibi). Bu sonuçlar burada belirtilen eşik değerlerine göre oluşturulmaktadır.
- 4) **Dikkat dağınıklığı tespit edicisi ayarları (attention tracker settings):** Opsiyonel bir ayardır. Sınıfta kaç öğrenci olduğu buradan girilebilmektedir. Burada belirtilen sayı ile

görüntü işleme ile bulunan yüz sayısı aynı değilse ve “Accept ‘No Detection’ as distracted” seçilmişse bulunamayan yüzler “Dikkatsiz” kabul edilmekte ve sınıfın toplam dikkat oranları buna göre hesaplanmaktadır.

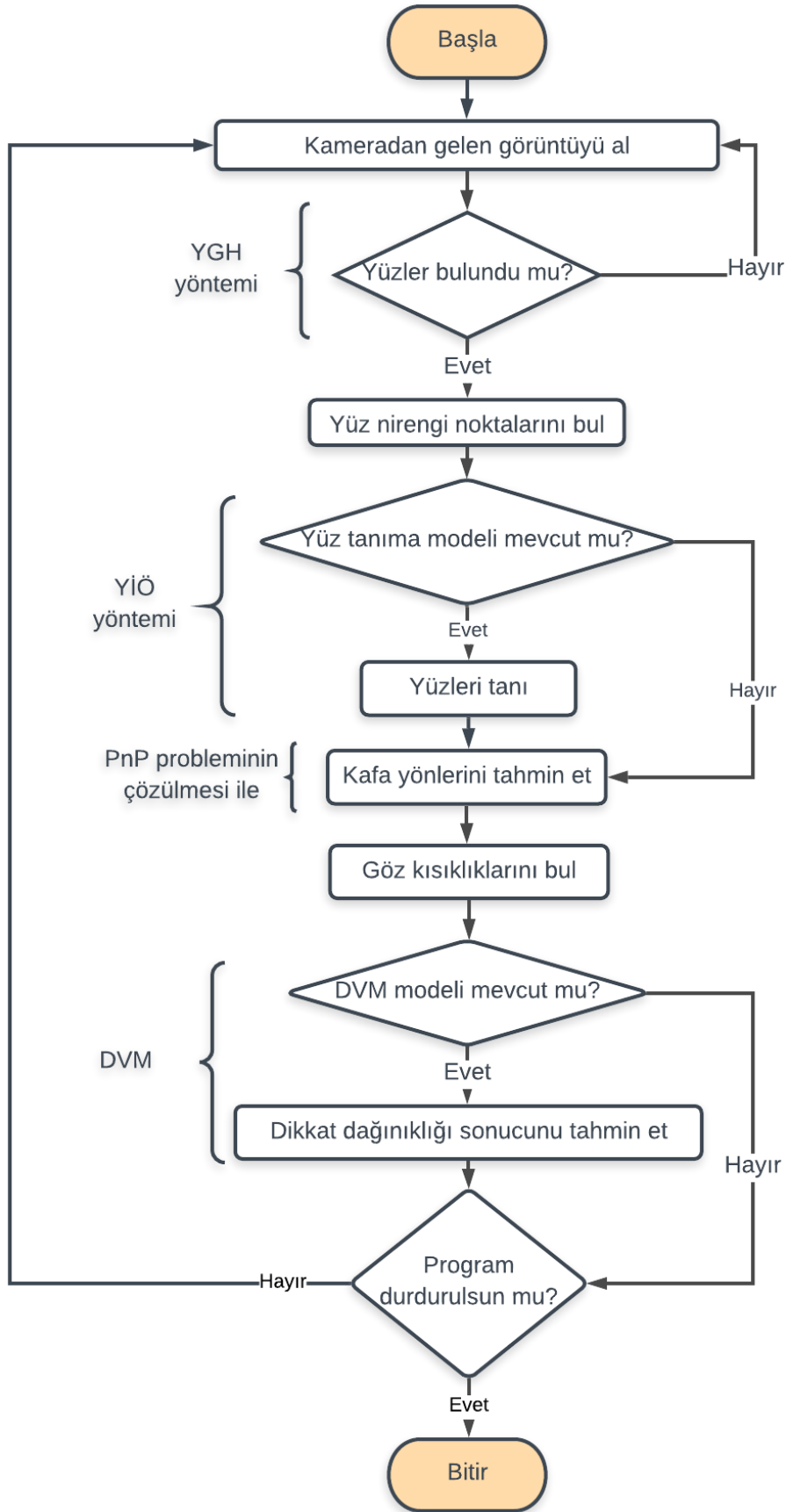
Belirtilen bu kısımlar Resim 4.5’de görülmektedir.

The image displays four panels of the DDTS software interface, each with a different colored border and a number indicating its position:

- Panel 1 (Red border):** Display Settings. Includes checkboxes for Camera, Landmarks, Head Pose Line, Rectangle Around Face, P-Y-R Values, P-Y-R Results, Attn. Tracking Time, and Attn. Tracking Results.
- Panel 2 (Blue border):** Performance Settings. Includes dropdown menus for Image Downsample Ratio (0.50) and Image Processing Ratio (1/4).
- Panel 3 (Purple border):** Head Pose Estimation Settings. Includes tabs for Pitch, Yaw, and Roll, and input fields for Min and Max values for Neutral, Up, and Down. An Apply button is also present.
- Panel 4 (Green border):** Attention Tracker Settings. Includes a Face Count input field (1) and a checkbox for "Accept 'No Detection' as distracted".

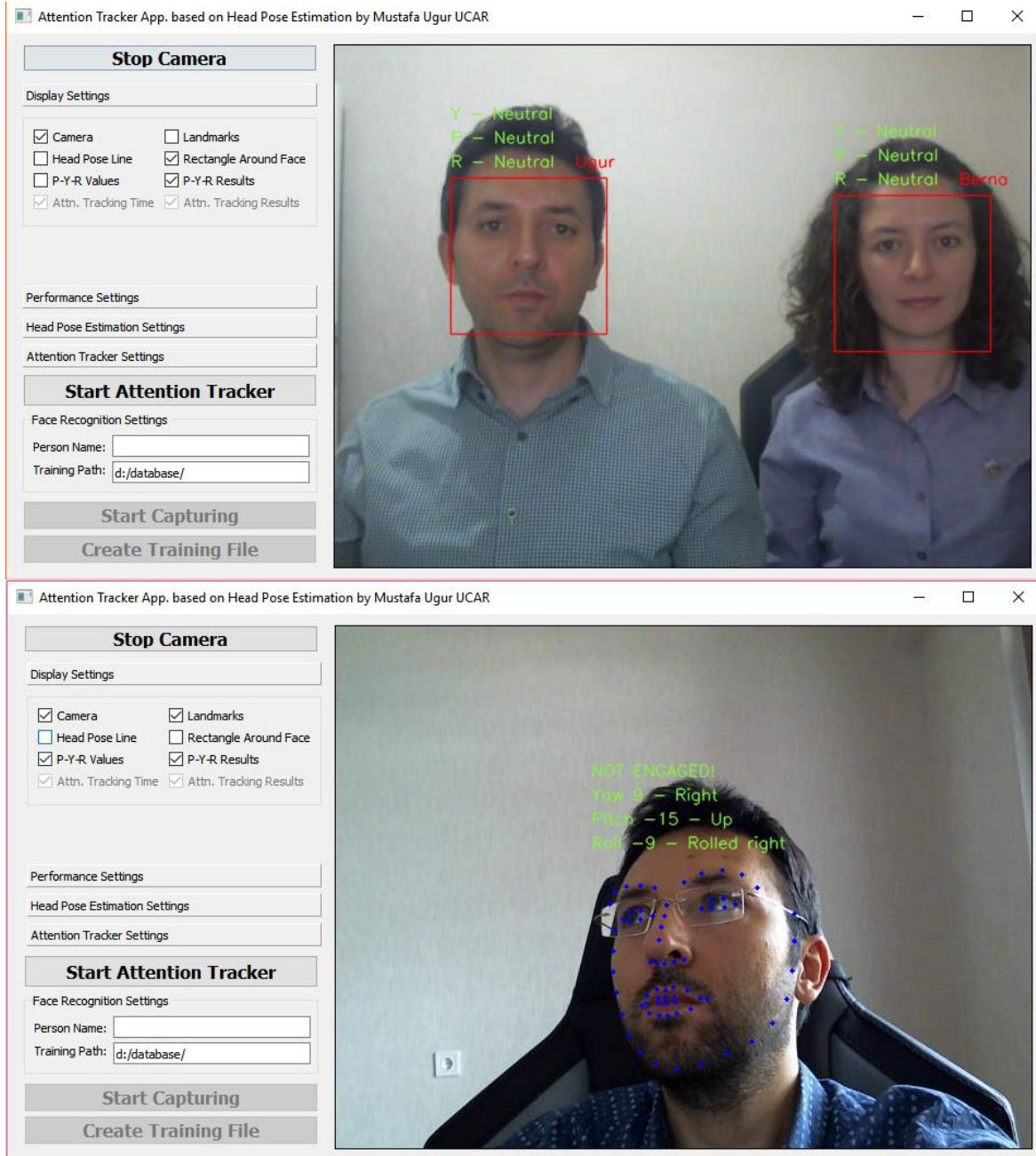
Resim 4.5. DDTS yazılımı ana bölümü - ayarlamalar

DDTS yazılımının ana bölümünün akış diyagramı Şekil 4.2’de görülmektedir.

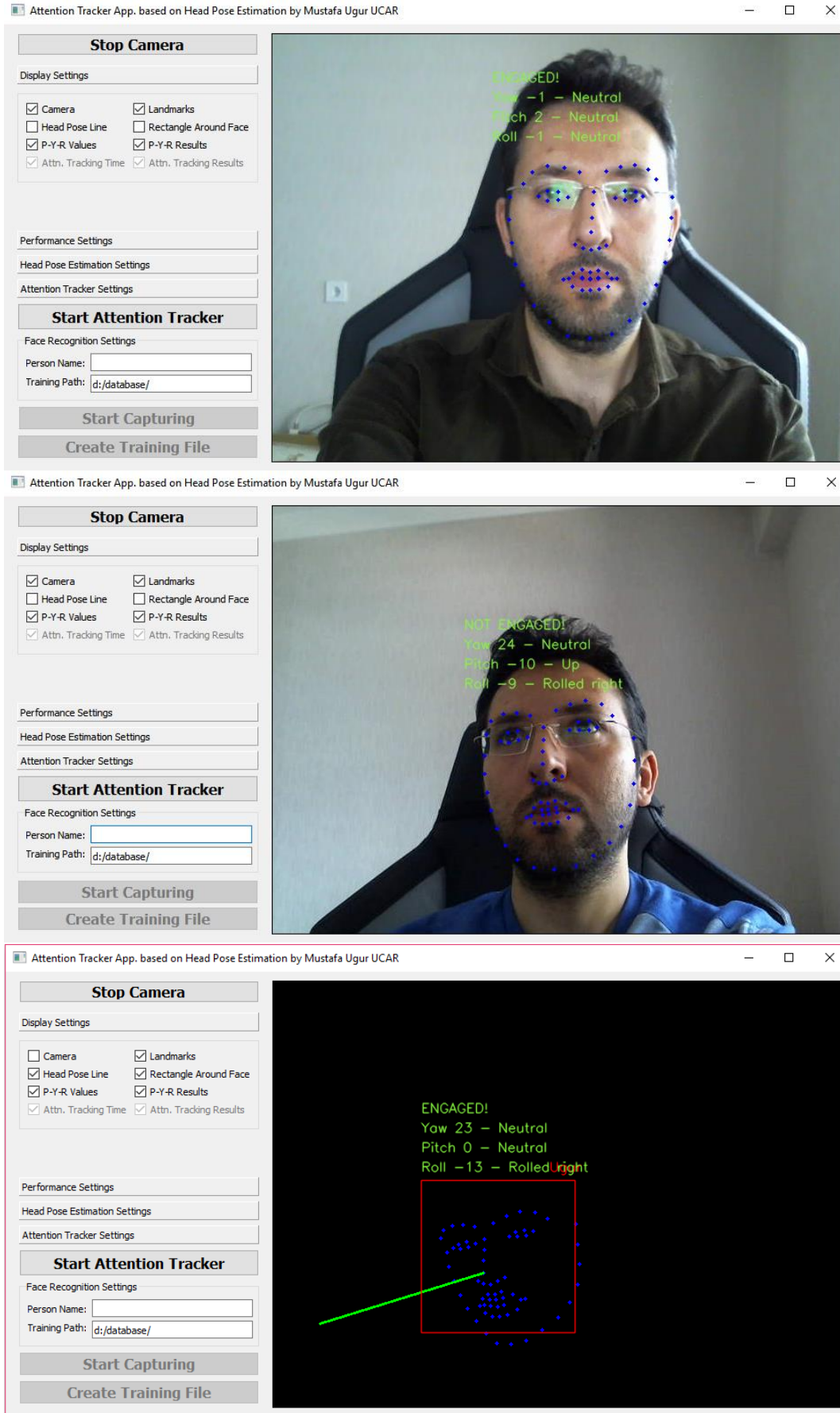


Şekil 4.2. DDTTS yazılımı yüz tanıma modeli oluşturma

DDTS ile ilgili yukarıda anlatılan ayarlamalar yapıldıktan sonra “Start Camera” butonuna basılarak görüntü yakalama başlatılır. Ayarlamalar görüntü yakalama esnasında da değiştirilebilmekte ve değiştirilen ayarların etkisi anlık olarak görülmektedir. Programın ana bölümü ile ilgili değişik ayarlamalara ilişkin çeşitli ekran görüntüleri Resim 4.6’da görülmektedir.



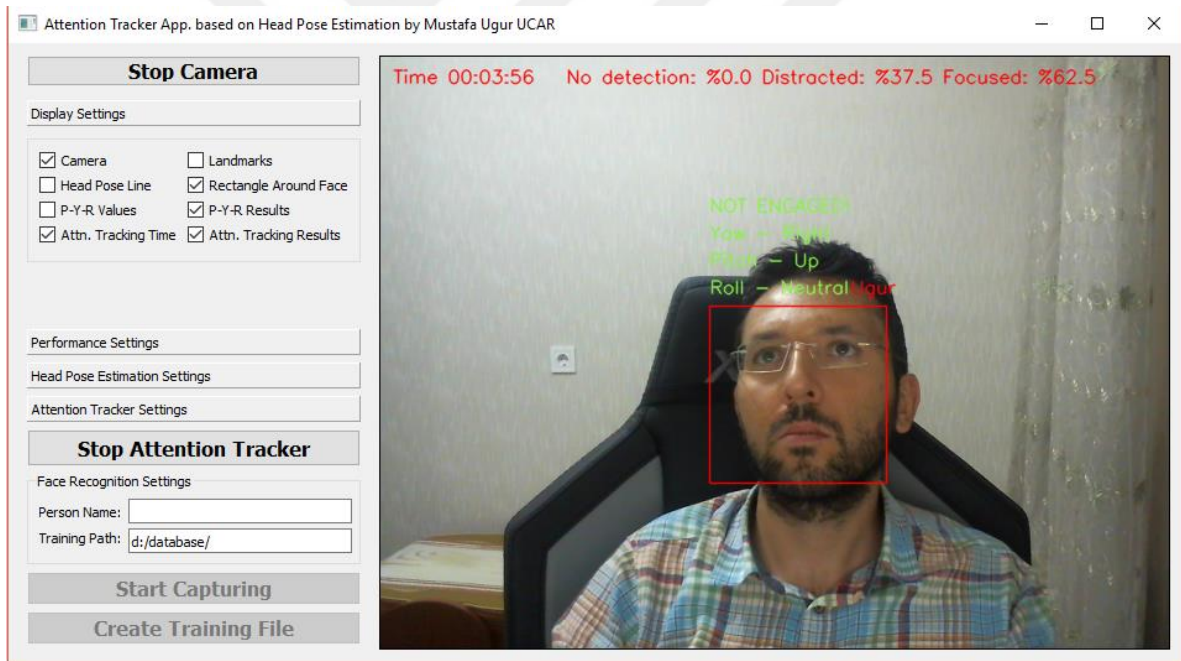
Resim 4.6. DDTS yazılımı ana bölümü – çeşitli ekran görüntüleri



Resim 4.6. (Devam) DDTs yazılımı ana bölümü – çeşitli ekran görüntüleri

Ekran görüntülerinde görünen “ENGAGED” (Dikkatli) ya “NOT ENGAGED” (Dikkatsiz) etiketi uygulamada kullandığımız DVM makine öğrenme algoritmasının her bir kare için gerçek zamanlı olarak üretmiş olduğu sonuçtur. DVM’nin neden seçildiği ve DVM’nin eğitim aşaması tezin “ARAŞTIRMA BULGULARI VE TARTIŞMA” bölümünde açıklanmıştır.

Dikkat dağınıklığı sonuçlarının anlık olarak takibinin yanı sıra belli zaman dilimindeki sonuçların kaydedilmesi ve raporlanabilmesi de mümkündür. “Start Attention Tracker” butonuna basıldıktan sonra uygulama her bir karede tespit edilen yüzleri tanır ve DVM’nin üretmiş olduğu sonuçları ilgili kişiler için kaydeder. Ayrıca tespit edilen yüzler için toplam dikkat dağınıklığı oranları da ekranda anlık olarak gösterilir. Bahsedilen toplam dikkat dağınıklığı sonuçları Resim 4.7’de görülmektedir.



Resim 4.7. DDTS yazılımı ana bölümü – anlık toplam dikkat dağınıklığı sonuçları

Resim 4.7’deki ekran görüntüsünü yorumlayacak olursak, 3 dakika 56 saniye boyunca kameranın tespit ettiği yüzlerin belirtilen süre içerisinde %62.5 oranında dikkatli oldukları görülmektedir.

Anlık raporlama işlemi “Stop Attention Tracker” butonuna basınca durdurulur ve dikkat dağınıklığı tespit sonuçları results.csv dosyasına kaydedilir. Bu dosyada tespit edilen kişi

isimleri ve takip edilen süre boyunca ilgili kişinin kaç karede “Dikkatli” kaç karede “Dikkatsiz” olduğu bilgileri bulunmaktadır. Yüz bulma işleminin gerçekleşmesi ama yüz tanıma modelinin olmaması ya da öğrencinin yüz tanıma modelinde yer almaması durumunda öğrenci ismi yerine “default” isminde bir satırda tespit edilemeyen kişiye ait dikkat dağınıklığı bilgisi gösterilmektedir. Örnek bir results.csv dosyasının görünümü Resim 4.8’de görülmektedir:

	A	B	C
1	PersonName	Engaged	NotEngaged
2	Ugur	4301	336
3	Berna	4214	423

Resim 4.8. DDTS yazılımı ana bölümü – dikkat dağınıklığı tespit sonuçları

5. ARAŞTIRMA BULGULARI ve TARTIŞMA

Tez çalışması kapsamında geliştirilen uygulamadan elde edilen bulguları 3 başlık altında inceleyebiliriz.

5.1. Kafa Yönü Tahmini ile İlgili Bulgular

Geliştirilen uygulamada kafa yönü tahminleri Bölüm 3.3.4'te açıklanan Perspective-n-Point probleminin çözülmesi ile elde edilmektedir. Matematiksel hesaplamalara dayalı bu yöntem herhangi bir makine öğrenmesi algoritması kullanmamaktadır.

Geliştirilen uygulamadaki kafa yönü tahmini ilgili kısmın başarısını tespit edebilmek amacıyla, UPNA Head Pose Database'de bulunan her bir videonun (veritabanında toplamda 120 video bulunmaktadır) her bir karesi için (her bir video 300 kareden oluşmaktadır), geliştirilen uygulama tarafından hesaplanan kafa yönü açıları tahmin edilerek kaydedilmiştir. Kaydedilen sonuçlar, UPNA Head Pose Database ile beraber gelen ve her bir videonun her bir karesi için sensörlerle ölçülmüş kafa yönü açıları (yunuslama, sapma ve yuvarlanma) ile kıyaslanmıştır. Kıyaslama sonuçları veritabanındaki kişilere göre gruplanarak yunuslama (pitch), sapma (yaw) ve yuvarlanma (roll) tahminlerindeki derece cinsinden ortalama mutlak sapmalar belirlenmiş ve bu sapmaların sırasıyla 1.34° , 4.97° ve 4.35° olarak gerçekleştiği gözlemlenmiştir. Detaylar Çizelge 5.1.'de gösterilmiştir.

Çizelge 5.1. Uygulamanın kafa yönü tahmini kısmının başarısı.

	Yunuslama	Sapma	Yuvarlanma
User_01	1.65	2.68	1.94
User_02	1.10	2.39	1.52
User_03	1.09	5.77	2.94
User_04	1.40	3.08	1.95
User_05	1.24	6.50	4.36
User_06	1.76	4.56	4.52
User_07	1.24	4.02	2.85
User_08	1.33	9.50	6.24
User_09	0.96	6.23	5.70
User_10	1.59	4.94	11.44
Ortalama	1.34	4.97	4.35

5.2. Yüz Tanıma ile İlgili Bulgular

Geliştirilen uygulamadaki yüz tanıma başarısının ölçülmesi amacıyla yine UPNA Head Pose Database kullanılmıştır. Bu amaçla veritabanındaki 10 kişiden her birinin video görüntülerinden çıkarılan 50 adet yüz fotoğrafına yüz hizalama, gri tona çevirme ve yeniden boyutlandırma işlemleri uygulanmış ve bu fotoğraflar yüz tanıma modelinin oluşturulması aşamasında kullanılmak üzere etiketlenerek kaydedilmiştir (Bkz. Şekil 4.1).

OpenCV kütüphanesinde hazır olarak gelmekte olan Yerel İkili Örüntüler (YİÖ), fisherfaces ve eigenfaces yüz tanıma algoritmaları, kaydedilen etiketlenmiş fotoğraflar ile eğitilerek her üç yöntem için yüz tanıma modelleri oluşturulmuştur. Sonrasında eğitilen YİÖ, fisherfaces ve eigenfaces yüz tanıma modelleri kullanılarak UPNA Head Pose Database’de bulunan her bir videonun her bir karesinde yüz tanıma gerçekleştirilmiş ve her bir kare için tahmin edilen etiketler kaydedilmiştir.

Tahmin edilen etiketler gerçek etiket değerleriyle karşılaştırıldığında YİÖ yönteminin %98.95 tanıma oranı ile en yüksek doğrulukta tanımayı gerçekleştirdiği gözlemlenmiştir. En iyi sonuçları elde ettiğimiz YİÖ yöntemi, geliştirilen uygulamada yüz tanıma amacıyla tercih edilmiştir. Fisherfaces ve eigenfaces yöntemleri kullanılarak gerçekleştirilen yüz tanıma işleminde de YİÖ’ye yakın ve yüksek doğrulukta sonuçlar elde edilmiştir. UPNA Head Pose Database’deki tüm kişiler için ortalama yüz tanıma doğruluk oranları fisherfaces yönteminde %98.69, eigenfaces yönteminde ise %94.57 olarak gerçekleşmiştir. Detaylar Çizelge 5.2.’de gösterilmiştir.

Çizelge 5.2. Uygulamanın yüz tanıma kısmının başarısı.

	YİÖ	Fisher	Eigen
User_01	98.69	99.86	90.13
User_02	100	100	99.25
User_03	96.36	97.58	90.22
User_04	98.97	96.16	95.58
User_05	100	100	96.22
User_06	99.52	99.75	96.52
User_07	99.08	100	95.77
User_08	99.80	99.88	97.38
User_09	99.83	96.36	94.11
User_10	97.30	97.33	90.55
Ortalama	98.95	98.69	94.57

5.3. Sistemin Geneli ile İlgili Bulgular

DDTS'nin genel başarısını, başka bir deyişle dikkat dağınıklığı tespitinin başarısını ölçmek adına iki ayrı test senaryosu gerçekleştirilmiştir.

İlk senaryoda makine öğrenme algoritmalarından bağımsız bir dikkat dağınıklığı tespiti yapıp yapılamayacağı üzerinde durulmuştur. Bu amaçla, geliştirilen uygulamada değişiklik yapılmış, yunuslama (pitch), sapma (yaw) ve yuvarlanma (roll) açıları için eşik değerleri belirlenerek o eşiklerden en az birisi geçildiğinde (üç açı değerinin en az bir tanesi nötrden farklıysa) öğrencinin kameraya bakmadığı, dolayısıyla da “dikkatsiz” olduğu kabul edilmiştir. Kullanılan web kamerasının görüş açısının yetersiz olması nedeniyle aynı anda 2 öğrenci ile 4 dakikalık bir test gerçekleştirilmiştir. Bu 4 dakikanın her bir dakikasında öğrenciler kendilerine verilen yönergeleri yerine getirmiştir. Örneğin 3. dakika boyunca Öğrenci1'in nötr bir şekilde durması istenirken Öğrenci2'nin başka yönlere bakması söylenmiştir. Öğrenci1 ve Öğrenci2 kendilerine verilen yönergeyi başlattıklarında “Start Attention Tracker” butonuna basılarak dikkat dağınıklığı tespiti başlatılmış ve 4. dakikanın bitimiyle de yönerge sonlandırılmıştır. Verilen yönergeler (solda) ve alınan sonuçlar Çizelge 5.3.'de gösterilmiştir.

Çizelge 5.3. DDTS'nin genel başarısı – 1. senaryo.

	YÖNERGE		SONUÇ	
	Öğrenci 1	Öğrenci 2	Öğrenci 1	Öğrenci 2
1. dakika	%100 dikkatli	%100 dikkatli	%96 dikkatli	%93 dikkatli
2. dakika	%50 dikkatli	%50 dikkatli	%50 dikkatli	%45 dikkatli
3. dakika	%100 dikkatli	%0 dikkatli	%88 dikkatli	%0 dikkatli
4. dakika	%70 dikkatli	%35 dikkatli	%66 dikkatli	%30 dikkatli

İkinci senaryoda ise makine öğrenme algoritmaları ile bir test gerçekleştirilmiştir. Bu amaçla UPNA Head Pose Database'deki 10 kişiye ait videolardan her bir kişi için 100 adet olmak üzere rastgele fotoğraflar kaydedilmiştir. Elde edilen 1000 fotoğraflık küme 5 kişiye dağıtılmış ve bu kümedeki her bir fotoğrafı “Dikkatli” ya da “Dikkatsiz” olarak sınıflandırmaları istenmiştir. Değerleyiciler bu sınıflandırmayı tamamladıktan sonra yaptıkları etiketleme işleminin tutarlı ve güvenilir olup olmadığı test edilmiştir. Bu amaçla değerlendiriciler arası güvenilirliği (inter-rater reliability) ölçen ve istatistiksel bir yaklaşım olan Fleiss' kappa katsayısı (Gwet, 2014) hesaplanmış ve 0.85 olarak bulunmuştur. Bu sonuç

değerleyiciler arasında yüksek bir tutarlılık olduğunu göstermiştir. Bu sonuç üzerine veri kümesindeki her bir fotoğraf 5 değerleyiciden çoğunluğunun görüşü (majority decision) ne ise o etiket ile etiketlenmiştir.

Oluşturulan veri kümesindeki “Dikkatli” ve “Dikkatsiz” olarak sınıflandırılan bazı fotoğraf örnekleri Şekil 5.1.’de görülmektedir.



Şekil 5.1. DDTS'nin genel başarısını test etmek amacıyla oluşturulan veri kümesinden örnek fotoğraflar (İlk satır: “Dikkatli”; İkinci satır: “Dikkatsiz”)

Etiketleme işlemi tamamlandıktan sonra bu 1000 fotoğraftan her birindeki kafa yönü açıları ve göz kısıklık oranları Bölüm 3’te anlatıldığı şekilde hesaplanarak kaydedilmiştir. Kaydedilen her bir fotoğraf için:

$$v = [yunuslama, sapma, yuvarlanma, göz_kapalılık_oranı, etiket]$$

şeklinde bir öznitelik vektörü çıkarılmış ve 1000 fotoğrafın yarısı eğitim, diğer yarısı da test kümesi olarak belirlenmiştir.

OpenCV içerisinde hazır olarak gelen Rastgele Orman (Random Forest), Karar Ağacı (Decision Tree), Destek Vektör Makineleri (Support Vector Machines) ve En Yakın Komşu (K Nearest Neighbor) makine öğrenme algoritmaları oluşturulan eğitim kümesi ile eğitilerek ilgili modeller yaratılmış ve bu modeller test kümesi üzerinde test edilmiştir.

Dikkat dağınıklığının tespitinde en başarılı sonuç %72.4 doğrulukla Destek Vektör Makineleri ile elde edilmiş ve geliştirilen uygulamanın dikkat dağınıklığı tespiti kısmında bu nedenle DVM'nin kullanılması tercih edilmiştir. Alınan sonuçlara ilişkin detaylar Çizelge 5.4.’de gösterilmiştir.

Çizelge 5.4. DDTS'nin genel başarısı – 2. senaryo (tahmin edilen Euler açıları ile).

	Doğruluk (%)
DVM	72.4
En Yakın Komşu	71.6
Rastgele Orman	70.6
Karar Ağacı	70.0

Ayrıca aynı senaryo için ikinci bir test de tahmin edilen Euler açıları yerine UPNA Head Pose Database ile birlikte gelen sensörlerle ölçülmüş Euler açıları kullanarak oluşturulan eğitim ve test kümesi üzerinden gerçekleştirilmiştir. Bu test sonucunda ise dikkat dağınıklığının tespitinde en başarılı sonuç %78.6 doğrulukla Rastgele Orman makine öğrenmesi algoritması ile elde edilmiştir. Alınan sonuçlara ilişkin detaylar Çizelge 5.5.'de gösterilmiştir.

Çizelge 5.5. DDTS'nin genel başarısı – 2. senaryo (ölçülen Euler açıları ile).

	Doğruluk (%)
DVM	76.4
En Yakın Komşu	77.0
Rastgele Orman	78.6
Karar Ağacı	77.8

6. SONUÇ ve ÖNERİLER

Tez çalışması kapsamında geliştirilen uygulama ile:

- UPNA Head Pose Database'de bulunan her biri 300 fotoğraf karesinden oluşan 120 video için görüntü işleme teknikleri kullanılarak kafa yönü açıları tahmin edilmiştir. Elde edilen yunuslama (pitch), sapma (yaw) ve yuvarlanma (roll) açıları sensörle hassas olarak ölçülmüş olan açı değerleri ile kıyaslanmıştır. Tahmin edilen yunuslama, sapma ve yuvarlanma açıları sırasıyla 1.34° , 4.97° ve 4.35° ortalama mutlak sapma değerleri ile elde edilmiştir.
- UPNA Head Pose Database'de bulunan 10 kişiye ait videolardan her bir kişi için 50 adet rastgele fotoğraf karesi çıkartılmış, bu fotoğraflar kişi isimleri ile ayrı klasörlere kaydedilmiş ve bu fotoğraflar kullanılarak yüz tanıma modeli oluşturulmuştur. Oluşturulan yüz tanıma modeli UPNA Head Pose Database'deki tüm videolar kullanılarak test edilerek YİÖ yöntemi ile %98.95 oranında bir yüz tanıma başarısı elde edilmiştir.
- UPNA Head Pose Database'deki videolardan çıkartılmış 1000 fotoğraf kullanılarak bir dikkat dağınıklığı veri kümesi oluşturulmuştur. Her bir fotoğraf için yunuslama, sapma, yuvarlanma ve göz kapalılık oranı hesaplanarak öznitelik vektörleri oluşturulmuştur. Veri kümesindeki her bir fotoğraf 5 kişi tarafından ayrı ayrı "Dikkatli" ya da "Dikkatsiz" olarak etiketlenmiştir. Etiketleme işleminin güvenilirliğini tespit etmek amacıyla Fleiss' kapa katsayısı 0.85 (yüksek güvenilirlik) olarak hesaplanmıştır. Veri kümesindeki her bir fotoğrafa etiket ataması yapılırken çoğunluk görüşü (majority decision) esas alınmıştır. Etiketlenmiş veri kümesinin yarısı makine öğrenme algoritmalarının eğitiminde, yarısı ise test edilmesinde kullanılmıştır. Yapılan testlerde DVM makine öğrenme algoritması %72.4 oranında bir dikkat dağınıklığı tespit başarısı göstermiştir.

Zaletelj ve Kosir (2017) tarafından yapılan çalışmada öğrencilerin dikkat dağınıklığının tespiti konusunda Kinect One sensörü kullanılarak elde edilen veriler ile eğitilen model dikkat dağınıklığının otomatik tespitinde %75.3'lük bir başarı göstermiştir. Monkaresi, Bosch, Calvo ve D'Mello (2017) tarafından yapılan çalışmada ise Kinect sensörünü kullanarak öğrenci dikkatinin tespitinde, aktivite sırasındaki raporlamalar kullanıldığında %75.8 ve aktivite sonrasındaki raporlamalar kullanıldığında ise %73.3 başarı elde edilmiştir.

Dikkat dađınıklığına tespitiyle ilgili yapılan bu alıřmaların sonuları dikkate alındığında, basit web kamerası kullanarak elde edilen sonuların Kinect sensörüyle yapılmıř olan alıřmalara yakın bir sonu verdiđi görülmektedir. Ayrıca geliřtirilen sistem öđrencinin tanınması ve dikkat dađınıklığı tespitinin öđrenci ile iliřkilendirilmesi yönünden daha kullanıřlıdır.

Bunun yanında geliřtirilen sistemde kullanılan web kamerasından dolayı görüř açısı ve mesafenin sınırlı kaldığı görülmüřtür. Sistemin sınıf ortamında daha başarılı sonular üretebilmesi için daha profesyonel, geniř açılı kameralar kullanılarak sınıf içi testlerin yapılması planlanmaktadır.

Dikkat dađınıklığına tespitini yapacak olan makine öđrenme algoritmalarının eğitimi ve testleri için 1000 fotođraftan oluřan etiketlenmiř bir veri kümesi oluřturulmuřtur. Bu veri kümesinin geniřletilmesi sonuları daha başarılı hale getirecektir.

Dikkat dađınıklığı sınıflandırıcılarının eğitilmesinde kullanılan öznitelik vektörü 5 elemandan oluřmaktadır. Bunlara ek olarak göz yönü, öđrencinin duygusal durumu, vücut yönü gibi özniteliklerin eklenmesinin dikkat dađınıklığı tespit başarısını artırmakta etkili olacağı öngörülmektedir.

Bu alıřmanın eğitimin her alanında otonom eğitim sistemlerinin geliřmesine ışık tutacağı düşünölmektedir.

KAYNAKLAR

- Afroze, S., & Hoque, M. M. (2016). Detection of Human's Focus of Attention using Head Pose. *1st International Conference on Information and Communication Technology 2016 (ICAICT 2016)*.
- Ahonen, T., Hadid, A., & Pietikäinen, M. (2006). Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), 2037-2041. doi:10.1109/tpami.2006.244
- Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The Perception of Where a Face or Television Portrait Is Looking. *The American Journal of Psychology*, 82(4), 474. doi:10.2307/1420441
- Apostoloff, N., & Zelinsky, A. (2003). Vision In and Out of Vehicles: Integrated Driver and Road Scene Monitoring. *Springer Tracts in Advanced Robotics Experimental Robotics VIII*, 634-643. doi:10.1007/3-540-36268-1_58
- Argyle, M., & Cook, M. (1976). Gaze and Mutual Gaze. *Cambridge University Press*. doi:10.2307/3032267
- Ariz, M., Bengoechea, J. J., Villanueva, A., & Cabeza, R. (2016). A novel 2D/3D database with automatic face annotation for head tracking and pose estimation. *Computer Vision and Image Understanding*, 148, 201-210. doi:10.1016/j.cviu.2015.04.009
- Ari, I., Uyar, A., & Akarun, L. (2008). Facial feature tracking and expression recognition for sign language. *2008 23rd International Symposium on Computer and Information Sciences*. doi:10.1109/iscis.2008.4717948
- Asteriadis, S., Tzouveli, P., Karpouzis, K., & Kollias, S. (2008). Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment. *Multimedia Tools and Applications*, 41(3), 469-493. doi:10.1007/s11042-008-0240-1
- Ayvaz, U., & Guruler, H. (2017). Real-time detection of students emotional states in the classroom. *25th Signal Processing and Communications Applications Conference (SIU)*. doi:10.1109/siu.2017.7960574
- Ba, S. O., & Odobez, J. M. (2009). Recognizing Visual Focus of Attention From Head Pose in Natural Meetings. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1), 16-33. doi:10.1109/tsmcb.2008.927274
- Ballantyne, M., Boyer, R. S., & Hines, L. M. (1996). Woody Bledsoe - His Life and Legacy. *AI Magazine*, 17, 7-20. doi:10.1609/aimag.v17i1.1207
- Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 711-720. doi:10.1109/34.598228
- Bertasius, G., Park, H. S., Yu, S. X., & Shi, J. (2017). Am I a Baller? Basketball Performance

- Assessment from First-Person Videos. *2017 IEEE International Conference on Computer Vision (ICCV)*. doi:10.1109/iccv.2017.239
- Bhattacharyya, D., Ranjan, R., Alisherov, F., & Choi, M. (2009). Biometric Authentication: A Review. *International Journal of U- and E- Service, Science and Technology (IJUNESST)*, 2(3), 13-28.
- Bidwell, J., & Fuchs, H. (2011). Classroom Analytics: Measuring Student Engagement with Automated Gaze Tracking. *Behavior Research Methods*, 49, 113. doi:10.13140/RG.2.1.4865.6242
- Bledsoe, W. W., & Chan, H. (1965). A Man-Machine Facial Recognition System: Some Preliminary Results. *Panoramic Research, Inc, Technical Report PRI 19A*, Palo Alto, California.
- Bledsoe, W. W. (1966a). Man-Machine Facial Recognition: Report on a Large-Scale Experiment. *Panoramic Research, Inc, Technical Report PRI 22*, Palo Alto, California.
- Bledsoe, W. W. (1966b). Some Results on Multicategory Pattern Recognition. *Journal of the Association for Computing Machinery*, 13(2), 304-316.
- Bock, S. W., Dicke, P., & Thier, P. (2008). How precise is gaze following in humans? *Vision Research*, 48(7), 946-957. doi:10.1016/j.visres.2008.01.011
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory (COLT'92)*, 144-152. doi:10.1145/130385.130401
- Bowyer, K. W. (2004). Face recognition technology: Security versus privacy. *IEEE Technology and Society Magazine*, 23(1), 9-19. doi:10.1109/mtas.2004.1273467
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 25, 120-125.
- Cengil, E., & Çınar, A. (2016). Görüntü sınıflandırma için yeni bir yaklaşım: Evrimsel Sinir Ağları. *European Journal of Technic*, 6(2), 96-103.
- Chervonenkis, A. Y. (2013). Early History of Support Vector Machines. *Empirical Inference*, 13-20. doi:10.1007/978-3-642-41136-6_3
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. doi:10.1007/bf00994018
- Cöster, J., & Ohlsson, M. (2015). Human Attention: The possibility of measuring human attention using OpenCV and the Viola-Jones face detection algorithm. BSc Thesis. *KTH Royal Institute of Technology, Sweden*.
- D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, 70(5),

377-398. doi:10.1016/j.ijhcs.2012.01.004

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455-470. doi:10.3758/bf03195475

Emery, N. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6), 581-604. doi:10.1016/s0149-7634(00)00025-7

Esas, M. Y. (2017). EOG testi ile dikkat eksikliğinin tanısı. Yüksek Lisans Tezi. *Erciyes Üniversitesi Fen Bilimleri Enstitüsü*, Kayseri.

Euler, L. (1776). *Novi Commentarii academiae scientiarum Petropolitanae*. 189-207.

Felsen, P., Agrawal, P., & Malik, J. (2017). What will Happen Next? Forecasting Player Moves in Sports Videos. *2017 IEEE International Conference on Computer Vision (ICCV)*, 3342-3351. doi:10.1109/iccv.2017.362

Fisher, R. A. (1936). The Use Of Multiple Measurements In Taxonomic Problems. *Annals of Eugenics*, 7(2), 179-188. doi:10.1111/j.1469-1809.1936.tb02137.x

Gao, X., Hou, X., Tang, J., & Cheng, H. (2003). Complete solution classification for the perspective-three-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8), 930-943. doi:10.1109/tpami.2003.1217599

Gibson, J. J., & Pick, A. D. (1963). Perception of Another Persons Looking Behavior. *The American Journal of Psychology*, 76(3), 386. doi:10.2307/1419779

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Gaithersburg, MD: Advanced Analytics, LLC.

Hacıoğlu, C. (2014). Derinlikli Öğrenme Kullanılarak Konuşmadan Uykululuk/Uykusuzluk Tespiti. Yüksek Lisans Tezi. *İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul.

Hesch, J. A., & Roumeliotis, S. I. (2011). A Direct Least-Squares (DLS) method for PnP. *2011 International Conference on Computer Vision (ICCV'11)*, 383-390. doi:10.1109/iccv.2011.6126266

Horand, R., Conio, B., Leboulleux, O., & Lacolle, B. (1989). An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics, and Image Processing*, 47(1), 33-44. doi:10.1016/0734-189x(89)90052-2

İnik, Ö., & Ülker, E. (2017). Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.

İnternet: Connor, N. (2018). Chinese school uses facial recognition to monitor student attention in class. URL: <https://www.telegraph.co.uk/news/2018/05/17/chinese-school-uses-facial-recognition-monitor-student-attention/>, Son Erişim Tarihi:

31.05.2019.

İnternet: dlib (2019). Official Web Page. URL: <http://dlib.net>, Son Erişim Tarihi: 02.06.2019.

İnternet: Farnsworth, B. (2017). Top 12 Eye Tracking Hardware Companies (Ranked). URL: <https://imotions.com/blog/top-eye-tracking-hardware-companies>. Son Erişim Tarihi: 21.05.2019.

İnternet: Mallick, S. (2016). Histogram of Oriented Gradients. URL: <https://www.learnopencv.com/histogram-of-oriented-gradients/>, Son Erişim Tarihi: 02.06.2019.

İnternet: Mallick, S. (2018). Support Vector Machines (SVM). URL: <https://www.learnopencv.com/support-vector-machines-svm/>, Son Erişim Tarihi: 03.06.2019.

İnternet: Morais, R. (2004). Seeing A Way To Safer Driving. *Forbes*. URL: https://www.forbes.com/2004/01/23/cz_rm_0123davosseeing.html. Son Erişim Tarihi: 21.05.2019.

İnternet: OpenCV (2019). Official web page. URL: <https://opencv.org/about/>, Son Erişim Tarihi: 02.06.2019.

İnternet: OpenCV - LBPH. Face Recognition with OpenCV. URL: https://docs.opencv.org/3.0-last-rst/modules/face/doc/facerec_tutorial.html, Son Erişim Tarihi: 03.06.2019.

İnternet: Prado, K. S. (2017). Face Recognition: Understanding LBPH Algorithm. URL: <https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b>, Son Erişim Tarihi: 03.06.2019.

İnternet: Slabaugh, G. G. (1999). Computing Euler angles from a rotation matrix. URL: <http://www.gregslabaugh.net/publications/euler.pdf>, Son Erişim Tarihi: 03.06.2019.

İnternet: Smith, S. (2017). Mobile Biometric Payment Volumes to Triple in 2017 to Nearly 2bn. URL: <https://www.juniperresearch.com/press/press-releases/mobile-biometric-payment-volumes-to-triple-in-2017>, Son Erişim Tarihi: 31.05.2019.

İnternet: Smith, S. (2018). Mobile Biometrics to Authenticate \$2 Trillion of Sales by 2023, Driven by Over 2500% Growth. URL: [https://www.juniperresearch.com/press/press-releases/mobile-biometrics-to-authenticate-\\$2-trillion-of-s](https://www.juniperresearch.com/press/press-releases/mobile-biometrics-to-authenticate-$2-trillion-of-s), Son Erişim Tarihi: 31.05.2019.

İnternet: Wikipedia - Image gradient. URL: https://en.wikipedia.org/wiki/Image_gradient/, Son Erişim Tarihi: 03.06.2019.

İnternet: Wikipedia-RRF URL: https://en.wikipedia.org/wiki/Rodrigues%27_rotation_formula, Son Erişim Tarihi: 03.06.2019.

- Jafri, R., & Arabnia, H. R. (2009). A Survey of Face Recognition Techniques. *Journal of Information Processing Systems*, 5(2), 41-68. doi:10.3745/jips.2009.5.2.041
- Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4-20. doi:10.1109/TCSVT.2003.818349
- Karol, S. (1963). *Zooloji Terimleri Sözlüğü*. Ankara, Türkiye: Türk Dil Kurumu Yayınları.
- Kekül, H., Bircan, H., & Arslan, H. (2018). Yüz tanıma uygulamalarında özyüzler ve yapay sinir ağlarının karşılaştırılması. *Uluslararası Yönetim Bilişim Sistemleri Ve Bilgisayar Bilimleri Dergisi*, 2(1), 51-59.
- Keller, C. G., & Gavrilu, D. M. (2014). Will the Pedestrian Cross? A Study on Pedestrian Path Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 15(2), 494-506. doi:10.1109/tits.2013.2280766
- King, D. E. (2009). Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research*, 10, 1755-1758.
- Kluttz, N. L., Mayes, B. R., West, R. W., & Kerby, D. S. (2009). The effect of head turn on the perception of gaze. *Vision Research*, 49(15), 1979-1993. doi:10.1016/j.visres.2009.05.013
- Krithika, L. B., & Lakshmi Priya, G. G. (2016). Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric. *Procedia Computer Science*, 85, 767-776. doi:10.1016/j.procs.2016.05.264
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90. doi:10.1145/3065386
- Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, 4(2), 50-59. doi:10.1016/s1364-6613(99)01436-9
- LeCun, Y. (1988). A theoretical framework for back-propagation. *Proceedings of the 1988 Connectionist Models Summer School, CMU, Pittsburg, PA*, 21-28.
- Lepetit, V., Moreno-Noguer, F., & Fua, P. (2009). EPnP: An Accurate O(n) Solution to the PnP Problem. *International Journal of Computer Vision*, 81(2), 155-166. doi:10.1007/s11263-008-0152-6
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2), 164-168. doi:10.1090/qam/10666
- Lock, J. C. (2016). Pose Error Estimation of a Quadcopter in the Outdoors. MSc Thesis. *Stellenbosch University, Cape Town, South Africa*.
- Lu, C. P., Hager, G. D., & Mjolsness, E. (2000). Fast and globally convergent pose estimation from video images. *IEEE Transactions on Pattern Analysis and Machine*

Intelligence, 22(6), 610-622. doi:10.1109/34.862199

- Marquardt, D. W. (1963). An Algorithm for Least Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2), 431-441. doi:10.1137/0111030
- Monkaresi, H., Bosch, N., Calvo, R. A., & Dmello, S. K. (2017). Automated Detection of Engagement Using Video-Based Estimation of Facial Expressions and Heart Rate. *IEEE Transactions on Affective Computing*, 8(1), 15-28. doi:10.1109/taffc.2016.2515084
- Moore, S., & Bowden, R. (2011). Local binary patterns for multi-view facial expression recognition. *Computer Vision and Image Understanding*, 115(4), 541-558. doi:10.1016/j.cviu.2010.12.001
- Moors, P., Verfaillie, K., Daems, T., Pomianowska, I., & Germeys, F. (2016). The Effect of Head Orientation on Perceived Gaze Direction: Revisiting Gibson and Pick (1963) and Cline (1967). *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.01191
- Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., & Goodman, M. (2005). An Overview of the 100-Car Naturalistic Study and Findings. *NHTSA Report, No.05-400*, Washington DC.
- Ojala, T., Pietikäinen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *Proceedings of 12th International Conference on Pattern Recognition (ICPR 1994)*, 1, 582-585. doi:10.1109/icpr.1994.576366
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), 51-59. doi:10.1016/0031-3203(95)00067-4
- Otsuka, Y., Mareschal, I., Calder, A. J., & Clifford, C. W. (2014). Dual-route model of the effect of head orientation on perceived gaze direction. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1425-1439. doi:10.1037/a0036151
- Ömur, S., & Aydoğdu, A. G. (2017). Eye tracking researches and new trends in the field of communication. *International Journal of Social Sciences and Education Research*, 3(4), 1296-1307. doi:10.24289/ijsser.321815
- Penate-Sanchez, A., Andrade-Cetto, J., & Moreno-Noguer, F. (2013). Exhaustive Linearization for Robust Camera Pose and Focal Length Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(10), 2387-2400. doi:10.1109/tpami.2013.36
- Raca, M. (2015). Camera-based estimation of student's attention in class. PhD Thesis. *École Polytechnique Fédérale De Lausanne*, Switzerland.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-

- propagating errors. *Nature*, 323(6088), 533-536. doi:10.1038/323533a0
- Schweighofer, G., & Pinz, A. (2006). Robust Pose Estimation from a Planar Target. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), 2024-2030. doi:10.1109/tpami.2006.252
- Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, 4(3), 519-524. doi:10.1364/josaa.4.000519
- Soukupová, T., & Cech, J. (2016). Real-Time Eye Blink Detection using Facial Landmarks. *21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia*.
- Stiefelhagen, R., & Zhu, J. (2002). Head orientation and gaze direction in meetings. *CHI 02 Extended Abstracts on Human Factors in Computing Systems - CHI 02*. doi:10.1145/506443.506634
- Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology*, 131, 451-469.
- Sümer, Ö, Goldberg, P., Sturmer, K., Seidel, T., Gerjets, P., Trautwein, U., & Kasneci, E. (2018). Teachers' Perception in the Classroom. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2315-2324.
- Tripathi, R. K., Jalal, A. S., & Agrawal, S. C. (2017). Suspicious human activity recognition: A review. *Artificial Intelligence Review*, 50(2), 283-339. doi:10.1007/s10462-017-9545-7
- Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. *Proceedings 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'91)*, 586-591. doi:10.1109/cvpr.1991.139758
- Ulusoy, B. (2015). Dikkat dağınıklığı üzerine bir beyin bilgisayar arayüzü tasarımı ve uygulaması. Yüksek Lisans Tezi. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya*.
- Uğur, M. E. (2018). Öğrenme Yönetim Sistemlerinde Derse Katılım Durumunun Yüz Tanıma Teknolojisi ile Belirlenmesi. Yüksek Lisans Tezi. *Dumlupınar Üniversitesi Fen Bilimleri Enstitüsü, Kütahya*.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer. doi:10.1007/978-1-4757-2440-0
- Vertegaal, R., Slagter, R., Veer, G. V., & Nijholt, A. (2001). Eye gaze patterns in conversations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 01*. doi:10.1145/365024.365119
- Wang, B., & Hoai, M. (2018). Predicting Body Movement and Recognizing Actions: An Integrated Framework for Mutual Benefits. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 341-348. doi:10.1109/fg.2018.00056

- Westin, A. F. (2006). How the Public Sees the Security-versus-Liberty Debate. In *Protecting what matters: Technology, security, and liberty since 9/11* (pp. 19-36). Washington: Brookings Institution.
- Whitehill, J., Serpell, Z., Lin, Y., Foster, A., & Movellan, J. R. (2014). The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions. *IEEE Transactions on Affective Computing*, 5(1), 86-98. doi:10.1109/taffc.2014.2316163
- Wu, Z., Singh, B., Davis, L. S., & Subrahmanian, V. S. (2018). Deception Detection in Videos. *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, 1695-1702.
- Wollaston, W. H. (1824). XIII. On the apparent direction of eyes in a portrait. (1824). *Philosophical Transactions of the Royal Society of London*, 114, 247-256. doi:10.1098/rstl.1824.0016
- Yampolskiy, R. V., & Govindaraju, V. (2008). Behavioural biometrics: A survey and classification. *International Journal of Biometrics*, 1(1), 81-113. doi:10.1504/ijbm.2008.018665
- Yücel, Z., Salah, A. A., Meriçli, Ç., & Meriçli, T. (2010). Kafa Duruşu Kestirimlerinden Bakış Yönünün Türetilmesi. *SIU 2010 - IEEE 18th Signal Processing and Communications Applications Conference*, 224-227.
- Zaletelj, J., & Košir, A. (2017). Predicting students' attention in the classroom from Kinect facial and body features. *EURASIP Journal on Image and Video Processing*, 2017(1), 80. doi:10.1186/s13640-017-0228-8
- Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. *Computer Vision – ECCV 2014 Lecture Notes in Computer Science*, 818-833. doi:10.1007/978-3-319-10590-1_53
- Zelinsky, A. (2006). Smart Cars: An Ideal Applications Platform for Robotics and Vision Technologies. *Colloquium on Robotics and Automation*, Napoli.
- Zhang, W., Cheng, B., & Lin, Y. (2012). Driver drowsiness recognition based on computer vision technology. *Tsinghua Science and Technology*, 17(3), 354-362. doi:10.1109/tst.2012.6216768

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : UÇAR, Mustafa Uğur
 Doğum tarihi ve yeri : 1983, Karaman
 e-mail : ugur.ucar@gmail.com



Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Yüksek lisans	İskenderun Teknik Üniversitesi / Elektrik-Elektronik Mühendisliği	2019
Lisans	Sakarya Üniversitesi / Bilgisayar Mühendisliği	2005
Lise	Karaman Yabancı Dil Ağırlıklı Lisesi	2001

İş Deneyimi

Yıl	Yer	Görev
2014-Hâlen	HAVELSAN	Yazılım Mühendisi
2007-2014	İSDEMİR	Yazılım Mühendisi
2007-2007	STC Elektronik	Yazılım Mühendisi

Yabancı Dil

İngilizce

Yayınlar

Uçar M. U. & Özdemir E. (2019). Detecting Student Engagement in e-learning Environment Based on Head Pose Estimation and Eye Aspect Ratio. *International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES'19)*, Proceeding Book, 119-123, Alanya, Turkey.

DİZİN

C

Convolutional Neural Networks ·
22

D

DDTS · ix, x, xi, xii, 47, 48, 49,
50, 51, 52, 53, 54, 55, 56, 57,
58, 61, 62, 63

Derin öğrenme · 21, 28, 68

Destek Vektör Makineleri · xii,
28, 38, 62

Dikkat Dağılımlığı Tespit Sistemi
· xii, 47

Dlib · iv, v, 27, 28, 70

DVM · iv, x, xii, 38, 39, 40, 41,
57, 62, 63

E

Eigenfaces · 20, 35, 66

En Yakın Komşu · iv, 62, 63

Euler açıları · ix, x, 43, 44, 63

Eye Aspect Ratio · 44, 74

F

face detection · 1, 19, 67

face recognition · v, 1, 19

Fisherfaces · 21, 35, 60, 66

Fleiss' kappa · 61, 64

G

görüntü işleme · iv, 1, 2, 3, 4, 8,
9, 10, 11, 22, 26, 27, 28, 30,
46, 52, 53

H

head pose estimation · v, 1, 8, 52

Histogram of Oriented Gradients
· 29, 32, 33, 34, 69

I

ILSVRC · x, xii, 22, 23

K

Karar Ağacı · iv, 62, 63

L

Local Binary Patterns · v, 9, 21,
35, 66

O

OpenCV · iv, v, 10, 12, 25, 26,
27, 28, 35, 42, 43, 44, 60, 62,
67, 69

P

Perspective-n-Point · iv, v, 28,
42, 59

pitch · iv, v, 43, 46, 52, 59, 61

R

Rastgele Orman · iv, 62, 63

roll · iv, v, 43, 46, 52, 59, 61

S

Support Vector Machines · 38,
41, 62, 67, 69

SVM · v, 38, 41, 69

U

UPNA Head Pose Database · iv,
v, xi, 45, 46, 59, 60, 61, 63

Y

yaw · iv, v, 9, 43, 46, 52, 59, 61

Yerel İkili Örüntüler · iv, xii, 9,
21, 28, 35, 60

Yönlendirilmiş Gradyanların
Histogramı · 28, 29

yüz bulma · 1, 28, 52

yüz tanıma · iv, ix, x, xi, 1, 4, 17,
18, 19, 20, 21, 26, 28, 36, 45,
48, 49, 50, 51, 52, 54, 58, 60



TEKNOVERSİTE



teknoversite **AYRICALIĞINDASINIZ**

İSTE

