T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜZ İFADELERİNİN GÖRÜNÜM TABANLI VE DİNAMİK ÖZELLİKLERİ KULLANILARAK OLUMSUZ KOŞULLAR ALTINDA HİBRİT YÜZ TANIMA

Murat TAŞKIRAN

DOKTORA TEZİ
Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Elektronik Programı

Danışman
Doç. Dr. Nihan KAHRAMAN

Eş-Danışman
Prof. Dr. Çiğdem EROĞLU ERDEM

Ocak, 2022

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜZ İFADELERİNİN GÖRÜNÜM TABANLI VE DİNAMİK ÖZELLİKLERİ KULLANILARAK OLUMSUZ KOŞULLAR ALTINDA HİBRİT YÜZ TANIMA

Murat TAŞKIRAN tarafından hazırlanan tez çalışması 18.01.2022 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Elektronik Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Doç. Dr. Nihan KAHRAMAN
Yildiz Teknik Üniversitesi
Danışman

Prof. Dr. Çiğdem EROĞLU ERDEM
Marmara Üniversitesi
Eş-Danışman

Jüri Üyeleri

Doç. Dr. Nihan KAHRAMAN, Danışman
Yildiz Teknik Üniversitesi _____

Prof. Dr. Tülay YILDIRIM, Üye
Yildiz Teknik Üniversitesi _____

Prof. Dr. Zümray DOKUR ÖLMEZ, Üye
İstanbul Teknik Üniversitesi _____

Doç. Dr. Umut Engin AYTEN, Üye
Yildiz Teknik Üniversitesi _____

Doç. Dr. Peyman MAHOUTİ, Üye
İstanbul Üniversitesi-Cerrahpaşa _____

Danışmanım Doç. Dr. Nihan KAHRAMAN sorumluluğunda tarafımca hazırlanan Yüz İfadelerinin Görünüm Tabanlı ve Dinamik Özellikleri Kullanılarak Olumsuz Koşullar Altında Hibrit Yüz Tanıma başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Murat TAŞKIRAN

İmza

Bu çalışma Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından EEAG-116E088 proje numarası ile desteklenmiştir.

Aileme ve dostlarıma
ithaf edilmiştir.

TEŞEKKÜR

Tezin başından sonuna kadar sağladığı değerli katkılarından dolayı öncelikle kıymetli danışmanım Doç. Dr. Nihan Kahraman'a ve değerli tez eş danışmanım Prof. Dr. Çiğdem Eroğlu Erdem'e, Tez İzleme Jürilerim Prof. Dr. Tülay Yıldırım ve Prof. Dr. Zümray Dokur Ölmez'e, tezin ilerleme aşamasında yaptıkları yönlendirmeler ve manevi desteklerinden dolayı çok teşekkür ederim. Tezin ilk aşamalarındaki katkılarından dolayı Mehmet Kıllıoğlu'na da teşekkürü bir borç bilirim. Lisansüstü eğitimim boyunca desteklerinden dolayı çalışma arkadaşım Arş. Gör. Hatice Vildan Düdükçü'ye de teşekkür ederim. Ayrıca tüm hayatım boyunca sonsuz desteğiyle her zaman bana güç veren babam Burhan Taşkıran ve annem Meryem Taşkıran'a ve ablam Tuğba Taşkıran'a teşekkür ederim. Bu doktora derecesi gözümde onlara aittir.

Son olarak özellikle lisansüstü eğitimim boyunca en büyük yardımcım olan, maddi manevi hiçbir desteğini esirgemeyen değerli arkadaşım, saygıdeğer meslektaşım ve sevgili eşim Zehra Gülru Çam Taşkıran'a ve tezimin son senesinde bana en büyük desteği sağlayan kızım Melisa Taşkıran'a sonsuz teşekkürlerimi sunarım.

Murat TAŞKIRAN

İÇİNDEKİLER

SİMGE LİSTESİ viii

KISALTMA LİSTESİ ix

ŞEKİL LİSTESİ xi

TABLO LİSTESİ xiii

ÖZET xv

ABSTRACT xvii

1 GİRİŞ 1

1.1 Literatür Özeti . . . 3

1.1.1 Yüz Tanıma Sistemlerinin Taksonomisi . . . 4

1.1.2 Yüz Tanıma Sistemlerinin Temel Adımları . . . 6

1.1.3 Veri Kümeleri . . . 9

1.1.4 Değerlendirme Ölçütleri . . . 9

1.1.5 Yüz Yanıltma Saldırıları . . . 16

1.1.6 Derin Öğrenme Temelli Yüz Tanıma Sistemleri . . . 17

1.2 Tezin Amacı . . . 21

1.3 Hipotez . . . 22

2 YÜZ HAREKETLERİNDEN YUMUŞAK BİYOMETRİK VERİ ÇIKARILMASI 23

2.1 Yüze Ait Dinamik Öznitelikler Kullanılarak Kimlik Bilgisi Doğrulama . . 24

2.1.1 Yüz Nirengi Noktalarının Tespit Edilmesi . . . 25

2.1.2 Yüze Ait Dinamik Özniteliklerin Çıkarılması. . . . 26

2.1.3 UvA-Nemo Veri Kümesi Üzerinde Yapılan Yüz Doğrulama Deneyleri . . . 28

3 YÜZ TANIMA İÇİN HİBRİT BİR YÖNTEM 35

3.1 Hibrit Yüz Tanıma Sistemi . . . 35

3.1.1 Yüz Tespiti . . . 35

3.1.2 Yüz Nirengi Noktalarının Tespit Edilmesi . . . 36

3.1.3 Görünüm-Temelli Özniteliklerin Elde Edilmesi 37
3.1.4 İstatistiksel Dinamik Yüz Özniteliklerinin Elde Edilmesi 40
3.1.5 Sınıflandırma 40
3.2 Hibrit Yüz Tanıma Sistemini Değerlendirmek için Gerçekleştirilen Deneysel Çalışmalar 41
3.2.1 Video Veri Kümeleri 41
3.2.2 Değerlendirme Metotları ve Test Edilen Bozulmalar 42
3.2.3 Video Veri Kümeleri Kullanılarak Gerçekleştirilen Deneysel Çalışmalar 44

4 SONUÇ VE ÖNERİLER **57**

KAYNAKÇA **61**

TEZDEN ÜRETİLMİŞ YAYINLAR **82**

SİMGE LİSTESİ

σ	Gauss Filtresi Standart Sapması
$(^{+})$	Gülümseme aktivitesinin başlangıç bölgesi
$(^{-})$	Gülümseme aktivitesinin bitiş bölgesi
$(^{a})$	Gülümseme aktivitesinin tepe bölgesi
l_i	İ'nci yüz nirengi noktası
ρ	İki yüz nirengi noktası arasındaki Öklid mesafesi
η	Kullanılan çerçeve sayısı
ω	Videonun kayıt edilmesi sırasındaki çerçeve hızı

KISALTMA LİSTESİ

ARAS	Aşırı Rastgele Ağaçlar Sınıflandırıcısı
AUROC	Alıcı Çalışma Özellikleri Altında Kalan Bölge
BBA	Bağımsız Bileşen Analizi
BTESA	Bölge Tabanlı Evrişimli Sinir Ağları
CMC	Kümülatif Eşleşme Karakteristiği
CoCo	Türdeş Kosinüs
ÇTBA	Çekirdek Temel Bileşen Analizi
DAA	Doğrusal Ayırıcı Analizi
DESA	Derin Evrişimli Sinir Ağları
DET	Tespit Hatası Değişimi
DSA	Derin Sinir Ağları
EER	Eşit Hata Oranı
ESA	Evrişimli Sinir Ağları
FAR	Yanlış Kabul Oranı
FMR	Yanlış Eşleşme Oranı
FMNR	Yanlış Eşleşmeme Oranı
FRR	Yanlış Reddetme Oranı
GAR	Gerçek Kabul Oranı
GMM	Gauss Karışım Modelleri
HOG	Yönlendirilmiş Gradyanların Histogramı
HYT	Hibrit Yüz Tanıma
k-NN	K-en yakın komşuluk
MMOD	Maksimum Marjin Nesne Tespiti

ÖÖD	Ölçekle-değişmeyen Özellik Dönüşümü
ROC	Alıcı Çalışma Özellikleri
SKS	Saniyedeki Kare Sayısı
SS	Standart Sapma
SSD	Tek Vuruş Çoklu-Kutu Dedektörü
STS	Seyrek Temsil Tabanlı Sınıflandırıcılar
TAD	Tek Atış Dedektörü
TAR	Doğru Kabul Oranı
TBA	Temel Bileşen Analizi
ÜÇA	Üretken Çekişmeli Ağlar
YDG	Yerel Olarak Doğrusal Gömme
YİD	Yerel İkili Desen
YOK	Yığılmış Otomatik Kodlayıcılar
ZSFA	Sıfır Vuruşlu Yüz Sahteciliği Önleme

ŞEKİL LİSTESİ

Şekil 1.1 Literatürdeki görüntü tabanlı ve video tabanlı yüz tanıma sistemlerinin taksonomisi. 5
Şekil 1.2 Yüz tanıma sistemlerinin ana adımları. Görüntüler UvA-NEMO veri kümesinden [2, 62] alınmıştır. 7
Şekil 1.3 Yüz tanıma için kullanılan veri kümeleri (1994-2005). (2B veri kümeleri mavi renkle, video veri kümeleri sarı renkle, 3B veri kümeleri yeşil renkle ve hiper-spektral/kızılötesi veri kümeleri kırmızı renkle gösterilir.) 10
Şekil 1.4 Yüz tanıma için kullanılan veri kümeleri (2006-2018). (2B veri kümeleri mavi renkle, video veri kümeleri sarı renkle, 3B veri kümeleri yeşil renkle ve hiper-spektral/kızılötesi veri kümeleri kırmızı renkle gösterilir.) 10
Şekil 2.1 Yüz dinamiklerine dayalı yüz tanıma yönteminin adımları. 25
Şekil 2.2 Kazemi ve arkadaşları tarafından önerilen yöntem kullanılarak tespit edilen 68 nirengi noktası ile UvA-NEMO veri tabanından bir örnek video karesi. 26
Şekil 2.3 Gülüşün bölümlere ayrılmış başlangıç, tepe noktası ve bitiş aşamaları. Çizim, videodaki kare numarasına karşı normalleştirilmiş ağız uzunluğunu göstermektedir. 28
Şekil 2.4 UvA-NEMO veri kümesindeki bir gülümseme videosundan elde edilen örnek kareler. 30
Şekil 2.5 Yalnızca spontan videolar için FMR ve FNMR diyagramı. 31
Şekil 2.6 Yalnızca kasıtlı videolar için FMR ve FNMR diyagramı. 31
Şekil 2.7 Bütün (spontan ve kasıtlı) videolar için FMR ve FNMR diyagramı. . 32
Şekil 2.8 UvA-NEMO gülümseme veri kümesinde, 128 boyutlu dinamik öznitelik vektörleri kullanılarak gerçekleştirilen yüz doğrulama deneyleri için FMR-FNMR diyagramı. 34
Şekil 2.9 UvA-NEMO gülümseme veri kümesinde, 128 boyutlu dinamik öznitelik vektörleri kullanılarak gerçekleştirilen yüz doğrulama deneyleri için Tespit Hatası Değişimi (DET) eğrisi. 34
Şekil 3.1 Önerilen hibrit yüz tanıma sisteminin blok diyagramı. 36

Şekil 3.2 VGG-16 derin sinir ağı mimarisi. [264] 38
Şekil 3.3 ResNet-50 derin sinir ağı mimarisi. [217] 39
Şekil 3.4 Ortalama (0,0) ve σ=1 için Gauss dağılımı. 43
Şekil 3.5 Farklı görüntü bozulmalarına sahip UvA-NEMO veri tabanından alınan örnek yüz görüntülerinde tespit edilen yüz işaretleri gösterilmektedir. (Parametreler parantez içinde belirtilmiştir. Yukarıdan aşağıya ve soldan sağa: Toplamsal Aydınlatma (+125), Kontrast Değişimi [Aralık: 0.49-0.51], Gauss Bulanıklığı ($\sigma = 10$), Çarpımsal Aydınlatma ($\times$ 2.5), Gauss Gürültüsü (Var = 0.1), Tuz ve Biber Gürültüsü (Oran = 0.3) gösterilmektedir.) 47
Şekil 3.6 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) ve HYT sistemler için UvA-NEMO gülümseme veri kümesini kullanılarak elde elde edilen yüz tanıma doğruluğu (%) grafikleri. 50
Şekil 3.7 Beş farklı bozulma altında HYT-VGGFace 2 için UvA-NEMO gülümseme veri kümesini kullanılarak elde edilen Kümülatif Eşleşme Karakteristik (CMC) eğrisi. 51
Şekil 3.8 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) ve HYT sistemler için FEEDTUM veri kümesini kullanılarak elde elde edilen yüz tanıma doğruluğu (%) grafikleri. 55
Şekil 3.9 Beş farklı bozulma altında HYT-VGGFace 2 için FEEDTUM veri kümesini kullanılarak elde edilen Kümülatif Eşleşme Karakteristik (CMC) eğrisi. 56

TABLO LİSTESİ

Tablo 1.1 Yüz tanıma sistemlerinin uygulama alanları. 2

Tablo 1.2 1994-2002 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.) 11

Tablo 1.3 2003-2009 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.) 12

Tablo 1.4 2010-2018 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.) 13

Tablo 1.5 1997-2018 yılları arasında yüz tanıma için yayınlanan video tabanlı veri kümeleri. (Kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet)) 14

Tablo 2.1 Dinamik öznitelikleri çıkarmak için kullanılan yüz mesafeleri. (ρ iki yüz nirengi noktası arasındaki Öklid mesafesini, l_i ise i'nci yüz nirengi noktasını belirtmektedir.) 27
Tablo 2.2 İstatistiksel Yüz Dinamiği Öznitelikleri. 29
Tablo 3.1 Kullanılan video veri kümelerinin açıklaması. 42
Tablo 3.2 Çeşitli bozulmalar için yalnızca istatistiksel yüz dinamiklerini kullanılarak UvA-NEMO gülümseme veri kümesinden elde edilen yüz tanıma doğrulukları (%). 46
Tablo 3.3 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için UvA-NEMO gülümseme veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GB" Gauss Bulanıklığı'nı, "TA" Toplamsal Aydınlatmayı, "ÇA" Çarpımsal Aydınlatmayı belirtmektedir.) 48
Tablo 3.4 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için UvA-NEMO gülümseme veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GG" Gauss Gürültüsünü, "KD" Kontrast Değişimlerini ve "TBG" Tuz Biber Gürültüsünü belirtmektedir.) . . . 49
Tablo 3.5 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için FEEDTUM veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GB" Gauss Bulanıklığı'nı, "TA" Toplamsal Aydınlatmayı, "ÇA" Çarpımsal Aydınlatmayı belirtmektedir.) 53
Tablo 3.6 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için FEEDTUM veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GG" Gauss Gürültüsünü, "KD" Kontrast Değişimlerini ve "TBG" Tuz Biber Gürültüsünü belirtmektedir.) 54

ÖZET

Yüz İfadelerinin Görünüm Tabanlı ve Dinamik Özellikleri Kullanılarak Olumsuz Koşullar Altında Hibrit Yüz Tanıma

Murat TAŞKIRAN

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Doktora Tezi

Danışman: Doç. Dr. Nihan KAHRAMAN
Eş-Danışman: Prof. Dr. Çiğdem EROĞLU ERDEM

Biyometrik sistemler, bir bireyin benzersiz fiziksel veya davranışsal özelliklerini ölçmeyi ve analiz etmeyi amaçlamaktadır. Biyometrik sistemlerin temel özelliği, kendine özgü özelliklere sahip vücut yapılarının kullanılmasıdır. Literatürde fizyolojik özellikleri (parmak izi, iris, avuç içi izi, yüz vb.) kullanan biyometrik sistemler ile davranışsal özellikleri (imza, yürüme, konuşma kalıpları, yüz dinamikleri vb.) kullanan sistemler bulunmaktadır. Son zamanlarda yüz biyometrisi, genellikle kullanıcının işbirliğini gerektirmeyen ve kişisel özel alanı ihlal etmeden elde edilebildiği için en çok tercih edilen biyometrik verilerden biri olmuştur.

Son zamanlarda derin öğrenme tabanlı yüz tanıma yöntemleri, büyük veri tabanlarında dikkate değer tanıma başarımları elde etse de, performanslarının olumsuz koşullar altında (örn. şiddetli aydınlatma ve kontrast değişimleri; bulanıklık ve gürültü) düştüğü görülmüştür. Bu koşullar altında, yüz dinamikleri gibi yumuşak biyometrik özniteliklerin, görünüm temelli özniteliklerle birlikte kullanılması halinde performansı artırması beklenmektedir. Bu tezde, gülümseme ifadesi sırasında derin evrişimli sinir ağlarından çıkarılan görünüme dayalı özniteliklerin yanı sıra, yüzün nirengi noktalarının konumlarından çıkarılan istatistiksel yüz dinamiklerini kullanan yeni bir hibrit yüz tanıma sistemi önerilmektedir. Farklı parametrelere sahip önemli görüntü bozulmaları altında, literatürdeki en başarılı yöntemler arasında yer alan, önceden eğitilmiş üç farklı derin evrişimli sinir ağlarının (VGG-Face, VGG-Face 2 ve

ArcFace) performansları değerlendirilmiştir. Deneysel çalışmaların sonuçları, yalnızca DESA tabanlı önitelikleri kullanan yüz tanıma performansının olumsuz koşullar altında önemli ölçüde düşmesine rağmen, yüz dinamikleri özniteliklerinin, derin evrişimli sinir ağı tabanlı özeliklerle birlikte kullanılmasının performans kaybını telafi edebileceğini ve doğruluğu önemli ölçüde artırabileceğini göstermektedir. Önerilen sistem, çok çeşitli aydınlatma varyasyonlarına sahip, bulanık ve gürültülü görüntüler içerebilen yüz tanıma problemlerinde ve yüz tanıma yanıltma ataklarına karşı faydalı olabilecektir.

Anahtar Kelimeler: Yüz tanıma, yüze ait dinamikler, yumuşak biyometrikler, derin sinir ağları, hibrit yüz tanıma sistemi

ABSTRACT

Hybrid Face Recognition Under Adverse Conditions Using Appearance-based and Dynamic Features of Facial Expressions

Murat TAŞKIRAN

Department of Electronics and Communication Engineering
Doctor of Philosophy Thesis

Supervisor: Doç. Dr. Nihan KAHRAMAN
Co-supervisor: Prof. Dr. Çiğdem EROĞLU ERDEM

Biometric systems have the goal of measuring and analyzing the unique physical or behavioral characteristics of an individual. The main feature of biometric systems is the use of bodily structures with distinctive characteristics. In the literature, there are biometric systems that use physiological features (fingerprint, iris, palm print, face, etc.) as well as systems that use behavioral characteristics (signature, walking, speech patterns, facial dynamics, etc.). Recently, facial biometrics has been one of the most preferred biometric data since it generally does not require the cooperation of the user and can be obtained without violating the personal private space.

Although recent deep-learning-based face recognition methods obtain remarkable recognition accuracies on large databases, their performance has been shown to degrade under adverse conditions (e.g. severe illumination and contrast variations; blur and noise). Under such conditions, soft-biometric features such as facial dynamics are expected to increase the performance if they are used together with appearance-based features. In this thesis, we propose a novel hybrid face recognition, which uses appearance-based features extracted using deep convolutional networks and statistical facial dynamics features extracted from facial landmark positions during smile expression. We evaluated the performances of three different state-of-the-art pre-trained deep convolutional neural networks (VGG-Face, VGG-Face 2 and ArcFace) under a variety of severe image distortions with different parameters. The experimental results show that, although the face recognition performance using

only DCNN-based features drops significantly under adverse conditions, the utilization of facial dynamics features together with DCNN-based features can compensate for the performance loss and increase the accuracy significantly. The proposed system is considered to be useful against face recognition problems, which may include blurry and noisy images with a wide variety of lighting variations, and spoofing attacks.

Keywords: Face recognition, facial dynamics, soft biometrics, deep neural network, hybrid face recognition system

1
GİRİŞ

Teknolojik gelişmeler günlük hayatımızda pek çok işlemi elektronik cihazlarla yapmamıza olanak sağlamaktadır. Bunun sonucu olarak elektronik ortamlarda yapılan işlemlerin güvenliği önemli bir sorun haline gelmiştir. Bu sorunun çözümü için, bir kişinin statik fizyolojik özellikleri (parmak izi, iris, yüz, avuç içi vb.) veya literatürde yumuşak biyometri olarak adlandırılan dinamik davranış özellikleri (imza, yürüme şekli, konuşma kalıpları ve yüz dinamikleri) kullanan biyometrik sistemlerin kimlik belirleme ve tanıma işlemleri için kullanımı hız kazanmıştır [1].

Yüz biyometrisi, güvenlik ve kolluk kuvvetleri, sağlık, eğitim, pazarlama, finans, eğlence ve insan-bilgisayar etkileşimi gibi birçok uygulama alanına sahip temel biyometrik özelliklerden biri olmuştur. Tablo 1.1'de yüz tanıma sistemlerinin ana uygulama alanları ve bu alanlarla ilgili özel uygulamalar listelenmiştir.

İnsan yüzü, kimlik, yaş [2], cinsiyet [3], ırk ve insanın duyguları ve zihinsel durumları yansıtan yüz ifadeleri [4–7] hakkında bilgiler taşımaktadır. İnsan yüzü ve yüz davranışının analizi, psikoloji, sinirbilim ve mühendisliği içeren disiplinler arası bir araştırma alanıdır. Yüz biyometrik verisi, diğer bazı biyometrik verilerin aksine, kişinin işbirliğini zorunlu değildir ve göze çarpmayan bir şekilde biyometrik verinin elde edilmesi gerçekleştirilebilir, bu durumda yüz biyometri verisini özellikle gözetim uygulamaları için uygun hale getirmektedir. Bununla birlikte, yüz tanıma sistemleri, yüzün hem fiziksel (statik) özniteliklerini hem de dinamik özniteliklerini kullanabilir ve bu da yüzün davranışsal biyometri için uygun bir biyometrik veri olmasını sağlar.

Yüzlerin algılanması, insanlar tarafından başarıyla ve neredeyse hiç çaba harcamadan gerçekleştirilen bir işlemdir, ancak bilgisayarlar için yüzlerin algılanması aynı kolaylıkta bir iş değildir. İnsan görsel sistemi, bağlamsal bilgiyle ilişkili olarak yüzleri tanımak için yüzlerin statik ve dinamik özniteliklerinin işlenmesi için karmaşık sinir yollarını barındırır [8]. Psikoloji ve sinirbilim alanında yüz algısı ile ilgili farklı sorunları ele alan birçok çalışma vardır. Örneğin, karakteristik özelliklerin baskın olabildiği halde hem bütünsel hem de özellik tabanlı temsillerin kullanıldığı

Tablo 1.1 Yüz tanıma sistemlerinin uygulama alanları.

Uygulama Alanları	**Spesifik Uygulamalar**
Güvenlik, Sağlık	Bilgi Güvenliği, kullanıcı yetkilendirme, Elektronik cihazlara erişim İnsan robot etkileşimi, binalara erişim Gözetim videolarından yüzleri tanıma Sağlıkla ilgili uygulamalar: robotik asistanlar Akıllı evler, akıllı araçlar
Hukuki Yaptırım	Sınır izleme, yasa dışı olay analizi, şüphelilerin takibi Pasaportlar, ulusal kimlik kartları, sürücü belgeleri Göçmenlik
Eğlence, Eğitim, Pazarlama	Oyun, sanal gerçeklik Fotoğraf yönetimi Video analizi, videodan bilgi çıkarımı Çevrimiçi öğrenme, öğrenci takibi, kullanıcı katılımı Reklam kampanyaları, sosyal ağ denetimi

gösterilmiştir [9]. Bruce ve Young tarafından ileri sürülen hipotez, yüz algısı için birlikte çalışan birkaç bağımsız alt süreç olduğudur [10]. Bu hipoteze göre, yüzün basit fiziksel özniteliklerinden bağımsız süreçler sonucunda yaş, cinsiyet ve temel yüz ifadeleri gibi çeşitli öznitelikler elde edilir ve bu özniteliklerin kullanılması kişisel bir yüz modeli moderasyon yapısının oluşturulmasını sağlar. Bu kişisel yüz modeli yapısı kullanılarak farklı koşullar altında bile yüzün algılanması beyin tarafından sağlanmaktadır. Yüzlerin tanınmasında hangi özelliklerin (gözler, ağız, burun vb.) daha önemli olduğunu anlamaya çalışan çalışmalar da vardır [11]. Hill ve Bruce tarafından gerçekleştirilen çalışmada [12], yüzlerin tanınması için üstten aydınlatmanın önemli olduğu gösterilmiştir. Olumsuzlama, ters çevirme veya eşikleme gibi zorlayıcı koşullar altında bile hareket halinde gösterildiklerinde tanıdık yüzlerin daha kolay tanındığını gösteren çalışmalar da hali hazırda literatürde mevcuttur [9, 13].

Son dönemlerde yüz tanıma ile ilgili literatürde gerçekleştirilen çalışmalar, normal koşullar altında yüz tanımanın, özellikle yüz tanıma sistemlerinde derin öğrenme ağlarının kullanımı ile başarılı bir şekilde gerçekleştirilebildiğini göstermiştir. Derin öğrenme yöntemleri, kontrollü ortamlar altında hem yüz tanımlama hem de doğrulama görevleri için yüksek performans elde etmelerine rağmen, performanslarının olumsuz koşullarda (örneğin aydınlatma, kontrast ve gürültü değişimleri) [14, 15] önemli ölçüde azaldığı görülmüştür. Olumsuz koşullar litaratürde semantik düşmanca saldırılar olarakda adlandırılmaktadır [16]. Bu tezde, olumsuz koşullar altında yüz tanıma başarımı azalan derin öğrenmeye dayalı yüz tanıma sistemlerinin, yüz tanıma başarımını artırmak amacı ile yumuşak

biyometrilerin yüz tanıma sistemine yardımcı veriler olarak kullanılması ve bu kullanılan yumuşak biyometri veriler sayesinde yüz tanıma sisteminin başarımının artırılması amaçlanmaktadır.

Birinci bölümün devamındaki alt başlıklarda, ilk olarak literatürdeki yüz tanıma çalışmalarının gelişimi ve elde ettikleri yüz tanıma başarımı hakkında bilgiler verilecek, daha sonra tezin amacı ve bu amaca ulaşmak için önerilen hipotezden bahsedilecektir. Bölüm 2'de yüz nirengi noktalarından elde edilen davranışsal dinamik özniteliklerin kişinin kimliğine ait bilgi taşıyıp taşımadığı incelenecek ve elde edilen sonuçlar sunulacaktır. Bölüm 3'te, olumsuz koşullar altında derin öğrenme ağlarının nasıl bir yüz tanıma performansı gösterdiği ve yüz nirengi noktalarından elde edilen istatistiksel dinamik özniteliklerin, derin öğrenme ağlarından elde edilen görünüm temelli öznitelikleri ile füzyonu hakkında bilgiler verilecektir. Üçüncü bölümün sonunda ise tez içerisinde önerilen hibrit yüz tanıma sisteminin elde ettiği yüz tanıma başarımları ve derin öğrenme ağları ile önerilen yöntemin kıyaslaması hakkında bilgiler verilecektir. Dördüncü ve son bölümde ise tezde gerçekleştirilen çalışmalar sonucunda elde edilen bulgular yorumlanacak ve ileride yapılabilecek çalışmalar hakkında öneriler sunulacaktır.

1.1 Literatür Özeti

Yüz tanımanın tarihi 1950'lere ve 1960'lara kadar uzanmaktadır, ancak otomatik yüz tanıma üzerine yapılan araştırmaların 1970'lerde başlatıldığı düşünülmektedir [17]. İlk çalışmalarda, yüzün önemli bölgeleri arasındaki mesafelere dayanan özellikler kullanılmıştır [18]. Yüz tanıma ile ilgili araştırma çalışmaları, donanımdaki gelişmeler ve güvenlikle ilgili uygulamalardaki artan önemin ardından 1990'ların başından itibaren hızla gelişmiştir.

1990'ların başından bu yana görüntüye dayalı yüz tanıma tekniklerinin gelişimi, Wang ve Deng [19] tarafından kabaca dört ana kavramsal geliştirme aşamasına bölünmüştür; fakat, tam bir taksonomiyi yansıtmamakta, daha çok temel yöntemlerin tarihsel gelişimini yansıtmaktadır: i) Bütünsel veya görünüme dayalı yaklaşımlar, yüz bölgesini bir bütün olarak kullanır ve yüzü daha düşük boyutlu bir alt uzaya eşlemek için doğrusal veya doğrusal olmayan yöntemler kullanır [20, 21]. İlk başarılı yöntemlerden biri Turk ve Pentland [22] tarafından geliştirilmiştir ve Eigenfaces adıyla bilinmektedir. Doğrusal alt uzayları [23], çok yönlü öğrenmeyi [24, 25] ve seyrek temsilleri [26, 27] kullanan başka yaklaşımlar da vardır. ii) Yerel öznitelik tabanlı yüz tanıma yöntemleri 2000'lerden sonra popüler hale geldi ve Gabor özellikleri [28, 29] ve yerel ikili desenler (LBP) ve varyantlar [30–32] gibi yüzü tanımlamak için

el yapımı öznitelikleri kullanmaktadırlar. iii) Öğrenmeye dayalı yerel tanımlayıcıları [33, 34] kullanan yöntemler 2010'lardan sonra ortaya çıkmış ve sığ teknikler kullanarak diskriminant görüntü filtrelerini öğrenmektedirler. iv) Derin Öğrenme tabanlı yöntemler, 2012 yılında düzenlenen ImageNet yarışmasında [35] AlexNet'in büyük başarısından sonra popülerlik kazanmıştır ve yüz tanıma probleminin çözümü için yeni bir bakış açısı getirmiştir. Yüz tanıma sistemleri için benzeri görülmemiş bir kararlılık elde edilmiştir, bunun sonucu olarak derin öğrenme temelli yüz tanıma sistemlerinin performansları, kısıtlanmamış ortamlardan toplanan büyük ölçekli veri kümelerindeki resimleri tanımada insan yüz tanıma performasına yaklaşmıştır [36, 37].

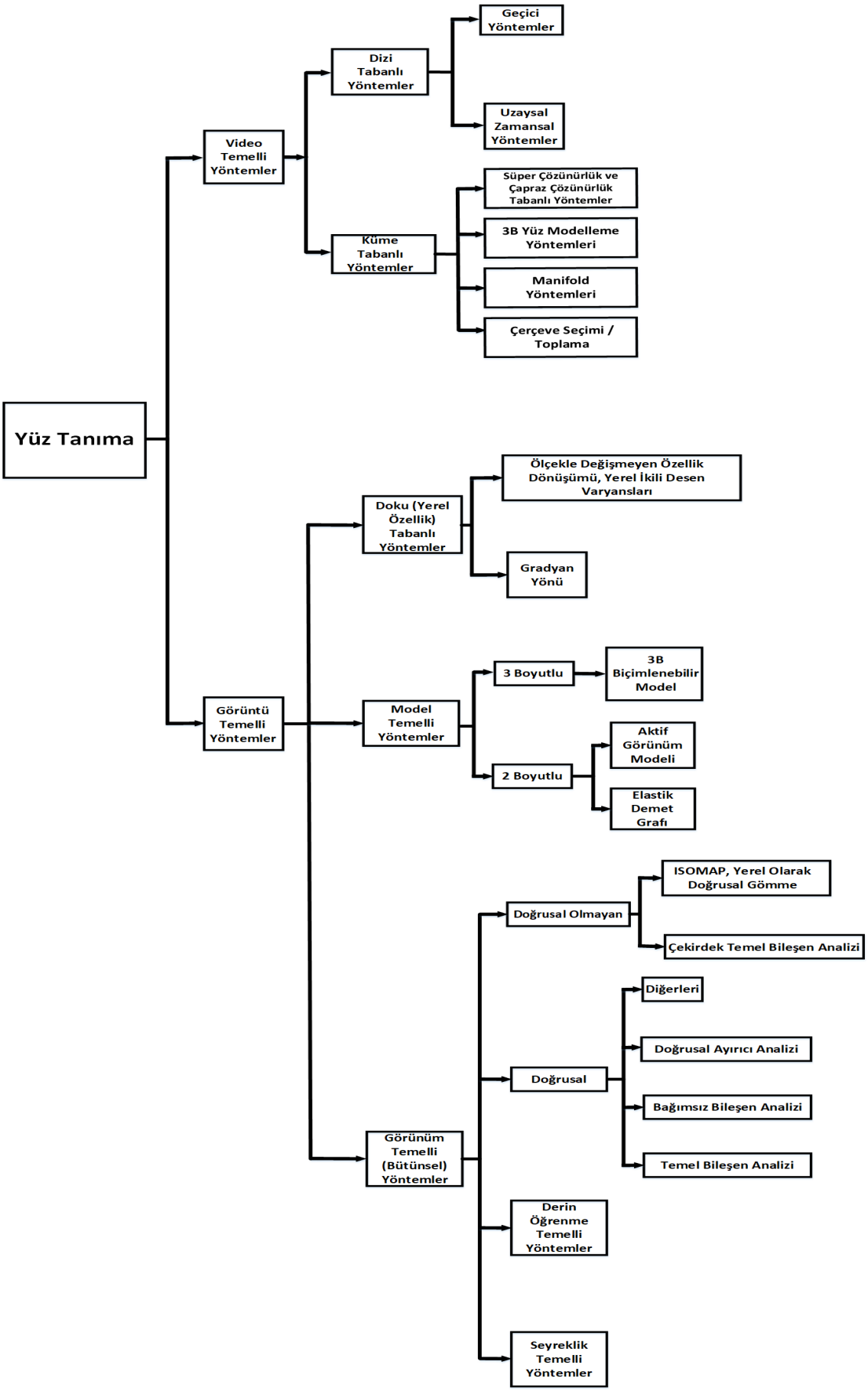
Literatürde yüz tanıma üzerine gerçekleştirilen çalışmaları özetleyen bir dizi derleme makalesi mevcuttur. Bu derleme makalelerinin ilkleri 1990'lı yıllarda yayınlanmıştır [38, 39]. Daha sonra, yüz tanıma sistemleri hakkındaki derleme makaleleri bazıları belirli bir yöne veya yönteme odaklanmış ve bu kısımla alakalı literatürdeki çalışmalar incelemişlerdir [17, 40–43].

- Poz [44, 45] veya ışıklandırma [46] ile değişmeyen yüz tanıma sistemleri,
- Videodan dinamik yüz tanıma [47–51],
- 3D ve kızılötesi modaliteleri kullanarak çok modlu yüz tanıma [52–54],
- Sunum saldırısı algılama (yüz yanıltma saldırıları önleme) yöntemleri [55, 56],
- Seyreklik tabanlı yüz tanıma yöntemleri [57],
- Derin öğrenme tabanlı yüz tanıma yöntemleri [19, 58–60].

Yüz tanıma işlemine bir tanımlama problemi veya bir doğrulama problemi olarak yaklaşılabilir. Yüz tanıma aynı zamanda 1:N eşleştirme problemi olarak da adlandırılır. Bilinmeyen yüz, bilinen kimlikler veri tabanındaki tüm yüzlerle karşılaştırılır ve tüm karşılaştırmalar sonucunda bir karar verilir. Kişinin veritabanında olduğu biliniyorsa görev kapalı küme, aksi takdirde açık küme olarak adlandırılır. Yüz doğrulama, 1:1 eşleştirme sorunu olarak bilinir. Sorgusu gerçekleştirilen yüzün kimliği, veri tabanında talep edilen kişinin kimliğine ait yüz verileriyle karşılaştırılarak doğrulanır veya reddedilir.

1.1.1 Yüz Tanıma Sistemlerinin Taksonomisi

Literatürdeki yüz tanıma sistemleri görüntü tabanlı ve video tabanlı yöntemler olarak iki ana gruba ayrılmaktadır. Görüntü tabanlı sistemler, fiziksel görünümü kullanarak



Şekil 1.1 Literatürdeki görüntü tabanlı ve video tabanlı yüz tanıma sistemlerinin taksonomisi.

bir kişiyi tanımaya çalışmaktayken, video tabanlı sistemler fiziksel görünümü ve zamanla görünümdeki değişiklikleri veya yüzün dinamiklerini kullanmaktadır. Yüz tanıma ile ilgili literatürün genel taksonomisi Şekil 1.1'de gösterilmektedir.

Görüntü tabanlı yüz tanıma yöntemleri üç ana grupta sınıflandırılabilir: i) Görünüme dayalı (veya bütünsel) yöntemler, ii) Model tabanlı yöntemler ve iii) Doku (yerel görünüm) tabanlı yöntemler [51, 61].

Video tabanlı yüz tanıma yöntemleri iki ana grupta sınıflandırılmaktadır: i) Küme tabanlı yöntemler ve ii) Dizi tabanlı yöntemler. Küme tabanlı yöntemler, bir videonun karelerini, karelerin zamansal sırasına dikkat etmeden bir görüntü koleksiyonu olarak ele almaktadır. Öte yandan, dizi tabanlı yöntemler, görüntüler arasındaki değişimleri kullanan yöntemlerdir. Dolayısıyla, yüzün zaman içinde değişim gösteren dinamikleri de kişinin tanınmasında rol oynamaktadır.

Literatürde yüz tanıma üzerine yapılan tüm çalışmaların net bir şekilde sınıflandırılması gerçekleştirilmesi zor bir işlemdir. Dolayısıyla, Şekil 1.1'de önerilen taksonomi, literatürde yüz tanıma için önerilen yöntemlerin kaba bir gruplamasıdır ve bazı gruplardaki algoritmaların birbiri ile örtüşen özelliklere sahip olmaları doğaldır.

1.1.2 Yüz Tanıma Sistemlerinin Temel Adımları

Yüz tanıma sistemleri geleneksel olarak altı ana aşamadan oluşmaktadır (bkz.Şekil 1.2):

(i) Yüzün giriş görüntüsü veya videosu alınır.

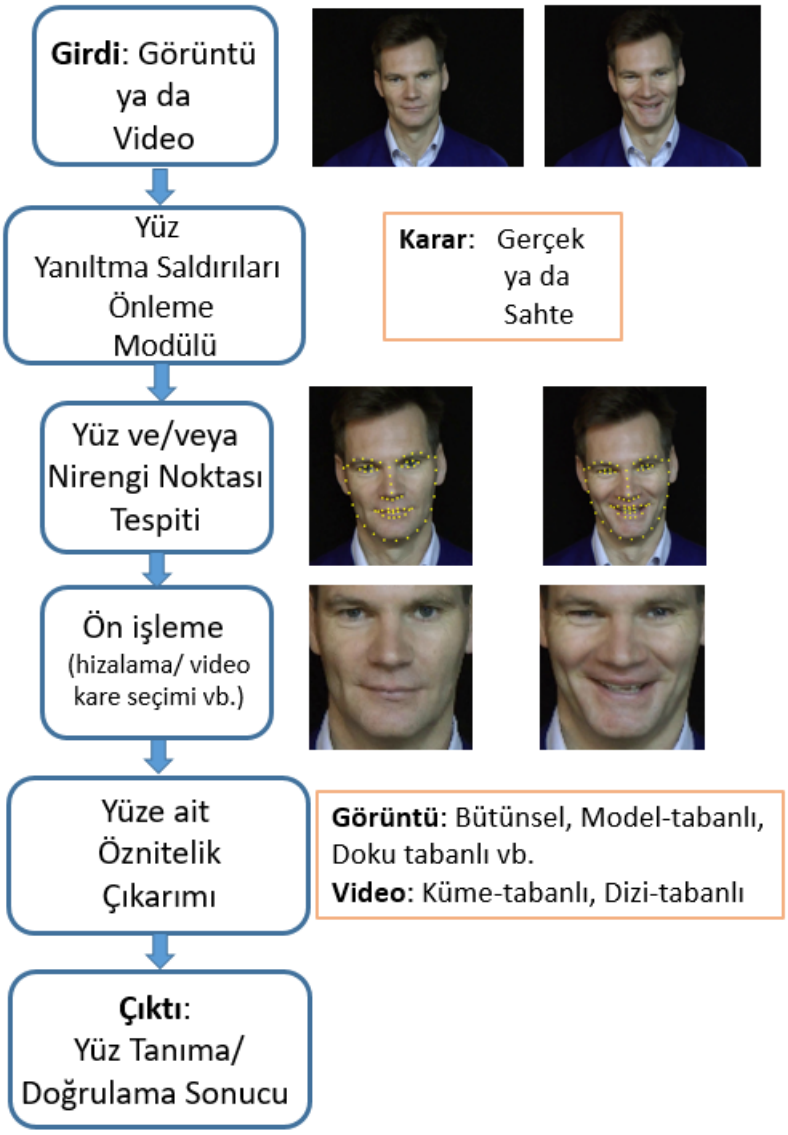
(ii) Yüz sahtekarlığı önleme modülü, sunum veya düşmanca saldırı tespiti (canlılık testleri vb. yoluyla) kullanarak sistemin güvenliğini sağlar.

(iii) Görüntüde veya her video karesinde yüz ve/veya yüz nirengi noktaları algılanır.

(iv) Görüntü veya video üzerinde, hizalama, video çerçeve seçimi, gürültü azaltma, kontrast iyileştirme veya benzer işlemlerden oluşabilecek ön işleme gerçekleştirilir.

(v) Görüntü veya videodan yüze ait öznitelik çıkarma. Görüntü tabanlı yöntemler bütünsel, model tabanlı veya doku tabanlı öznitelik çıkarma yaklaşımlarını kullanırken, video tabanlı yöntemler küme tabanlı veya sekans tabanlı yaklaşımları kullanır.

(vi) Yüz tanıma veya doğrulama işlemi gerçekleştirilir.



Şekil 1.2 Yüz tanıma sistemlerinin ana adımları. Görüntüler UvA-NEMO veri kümesinden [2, 62] alınmıştır.

Aşağıdaki alt başlıklarda, literatürdeki yüz algılama ve yüz nirengi noktalarının tespiti yöntemlerinin kısa bir incelemesini verilmektedir. Doğru ve etkili bir şekilde gerçekleştirilen yüz algılama ve yüz nirengi noktalarının tespiti algoritmaları sayesinde, yüz tanıma sistemlerinin doğruluğunu artırılabilmektedir.

1.1.2.1 Yüz Algılama

Yüz algılama tespiti, belirli bir görüntüdeki veya bir video karesindeki yüzün yerinin bulunarak, sınırlarının belirlenmesi işlemine verilen addır. Görüntülerde birden fazla yüz mevcutsa, yüz algılama algoritması tarafından yüzlerin hepsi algılanır. Yüz algılama algoritmaları, poz, aydınlatma ve ölçek farklılıklarına karşı dayanıklı olmalı ve arka planı mümkün olduğunca ortadan kaldırmalıdır [63].

Viola-Jones yüz algılayıcı [64], ön cepheden elde edilen yüz resimleri için iyi çalışan, yaygın olarak kullanılan bir yüz algılama yöntemidir. Haar benzeri özniteliklere dayanır ve gerçek zamanlı olarak çalışmaktadır. Viola-Jones yöntemi dışında, literatürde renk bilgisini kullanan başka yaklaşımlar da önerilmiştir [65–67].

Son zamanlarda, derin öğrenme tabanlı yüz algılama algoritmaları başarılı sonuçlar sağlamıştır [63, 68, 69]. Yakın zamanda önerilmiş olan bir yöntemde, bölge önerileri yaklaşımını kullanan, daha hızlı sonuç veren ve başlangıçta nesne tespiti için önerilmiş olan Bölge Tabanlı Evrişimli Sinir Ağları (BTESA) kullanılmıştır [70, 71]. Kayan pencere fikrini kullanan başka derin öğrenme tabanlı yüz algılama yöntemleri de vardır [72–74]. İlk olarak nesne algılama gerçekleştirmek için önerilen Tek Atış Dedektörü (TAD) [75], yüz algılama için de başarıyla kullanılmıştır [76, 77].

1.1.2.2 Yüz Nirengi Noktalarının Tespiti

Yüz tespit işlemi gerçekleştirildikten sonra, yüz hizalaması için yüzdeki yüz nirengi noktalarının (gözlerin köşeleri, kaşlar ve ağız, burun ucu vb.) tahmin edilme işlemi gerçekleştirilebilir. Yüzü kanonik bir konuma hizalamanın yüz tanıma için faydalı olduğu Bansal ve arkadaşaları tarafından gerçekleştirilen çalışmada gösterilmiştir [78]. Yüzdeki nirengi noktalarının gösterildiği örnek resim, regresyon ağaçları yaklaşımı [79] kullanılarak elde edilmiş ve Şekil 1.2'de sarı noktalarla gösterilmiştir.

2010'lu yılların başında yüz hizalama işlemlerinin yapılabilmesi için literatürde farklı yöntemler önerilmiş ve bu yöntemlerin yüksek performans gösterdiği yapılan deneysel çalışmalarda gösterilmiştir [80–82]. Yüzün nirengi noktalarının belirlenmesi üzerine gerçekleştiren çalışmaları özetleyen derleme makaleleri bulunmaktadır [83–90]. Wu ve Ji, yüz nirengi noktalarını algılama yöntemlerini bütünsel yöntemler,

Sınırlandırılmış Yerel Model (SYM) yöntemleri ve regresyon temelli yöntemler olarak üç kategoride sınıflandırmıştır [90]. Bir başka olası kategorizasyon ise, bu yöntemleri üretici yöntemler ve ayırt edici yöntemler olarak gruplandırmaktır [86].

Yüz nirengi noktalarının tespiti için geliştirilen algoritmaların performanslarını değerlendirmek için literatürde iki farklı metrik kullanılmaktadır:gerçek temele dayalı lokalizasyon hatası ve göreve yönelik performans [87]. Derin öğrenme tekniklerindeki en son gelişmeler sayesinde, yüz nirengi noktası çıkarma yöntemlerinin performansı, doğal ortamda elde edilen veri kümelerinde bile önemli ölçüde iyileştirilmiştir [83, 91, 92]. Çok görevli öğrenme için geliştirilmiş, yüz algılamayı ve nirengi nokta lokalizasyonunu diğer görevlerle birleştiren, tahmin ve cinsiyet tanımayı sağlayan yöntemlerde literatürde mevcuttur [93–95]. Son zamanlarda, derin sinir ağları kullanan mobil cihazlarda yüz takibi gerçekleştirilen çalışmalar bulunmaktadır [96].

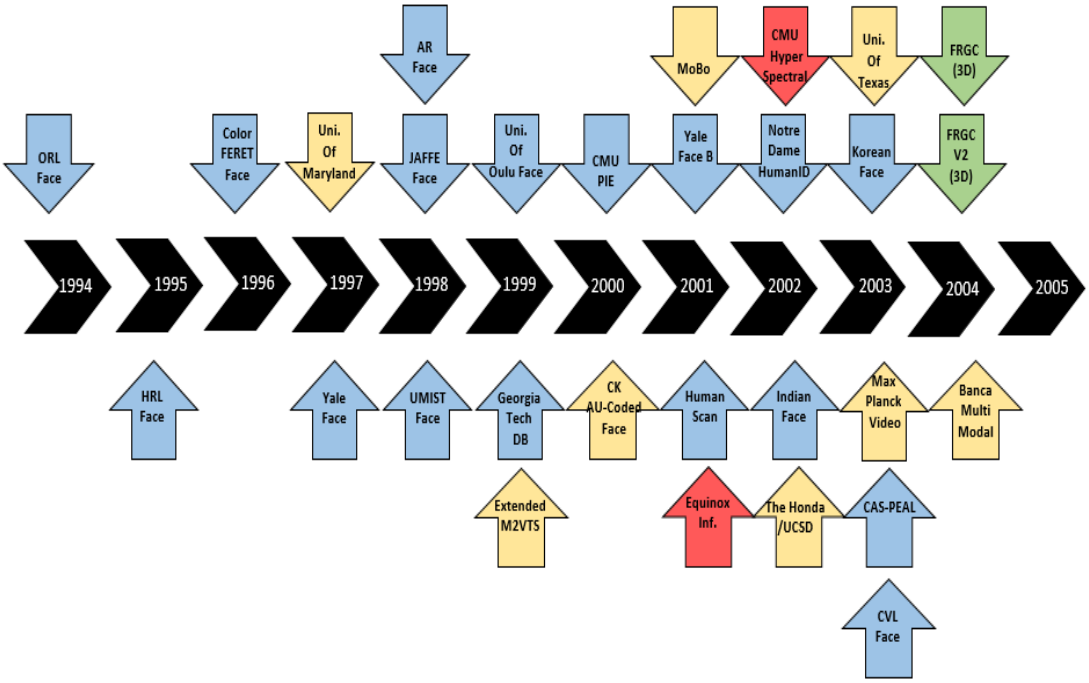
1.1.3 Veri Kümeleri

Yüz tanıma konusundaki ilk çalışmalar, kontrollü koşullar altında laboratuvarlarda kaydedilen oldukça küçük ölçekli veri kümelerine dayanıyordu. İlk görüntü tabanlı veri kümelerinden biri, 10 denekten 400 görüntü içeren ORL veri kümesidir [97]. Benzer şekilde, 1997'de yayınlanan ilk video tabanlı yüz veri kümelerinden biri [98], 40 denekten ve 70 videodan oluşmaktaydı. Son yıllarda, yüz tanıma veri kümeleri, kontrolsüz koşullarda kaydedilen milyonlarca görüntü veya on binlerce video ile büyük ölçekli hale gelmiştir.

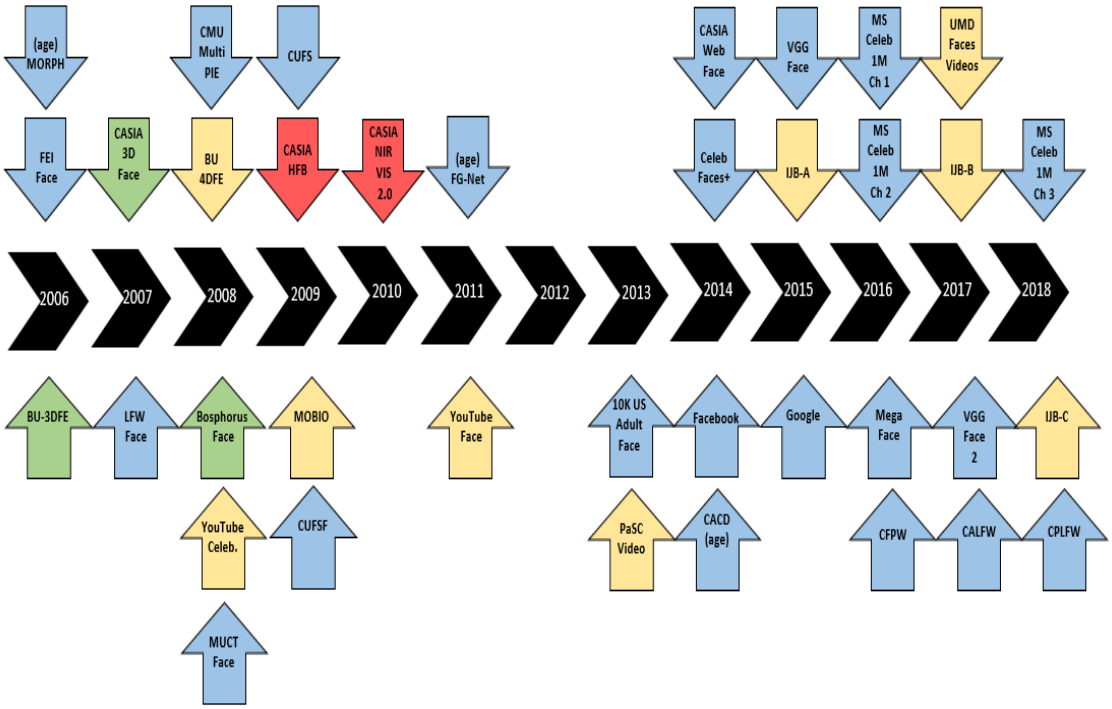
Yüz tanıma için kullanılan yüz veri kümeleri, görüntü tabanlı ve video tabanlı yüz veri kümeleri olarak iki ana başlık altında gruplandırılabilmektedir. Literatürdeki ana görüntü tabanlı ve video tabanlı yüz tanıma veri kümeleri sırasıyla Tablo 1.2, 1.3, 1.4 ve Tablo 1.5'te verilmiştir. Bununla birlikte, resim tabanlı, video ve 3 boyutlu yüz tanıma veri kümelerinin grafiksel zamansal gösterimi de Şekil 1.3 ve Şekil 1.4'te verilmektedir.

1.1.4 Değerlendirme Ölçütleri

Yüz tanıma sistemlerinin günlük hayatımızda kullanımının hızla artması ile bu sistemlerin başarımı çok önemli bir konu haline gelmiştir. Biyometrik sistemlerin performansını ölçmek için araştırmacılar tarafından bir takım değerlendirme ölçütleri önerilmiştir. Yüz tanıma, bir tanımlama veya doğrulama (kimlik doğrulama) yaklaşımı kullanılarak gerçekleştirilebilmektedir. Yüz doğrulama için yaygın olarak kullanılan değerlendirme ölçütleri ve çizelgeleri şunlardır [158, 159]:



Şekil 1.3 Yüz tanıma için kullanılan veri kümeleri (1994-2005). (2B veri kümeleri mavi renkle, video veri kümeleri sarı renkle, 3B veri kümeleri yeşil renkle ve hiper-spektral/kızılötesi veri kümeleri kırmızı renkle gösterilir.)



Şekil 1.4 Yüz tanıma için kullanılan veri kümeleri (2006-2018). (2B veri kümeleri mavi renkle, video veri kümeleri sarı renkle, 3B veri kümeleri yeşil renkle ve hiper-spektral/kızılötesi veri kümeleri kırmızı renkle gösterilir.)

Tablo 1.2 1994-2002 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.)

Veri Kümesi	**Yıl**	**Resim Boyutu, Renk**	**Denek Sayısı, Resim Sayısı**	**Özellikler**
ORL [97]	1994	92×110, H	10, 400	Bütün resimler ön cepheden elde edilmiş.
HRL [99]	1995	193×254, H	10, $\sim$ 800	AY: 77-84.
Color FERET [100]	1996	256×384, E	1199, 14.051	Yİ: 2, AY: 2, BP: 9-20, KS: 2.
Yale [101]	1997	320×243, H	15, 165	Gözlüklü ve gözlüksüz, AY, Yİ
JAFFE [102]	1998	256×256, H	10, 213	Yİ: 7
UMIST [103]	1998	220×220, H	20, 564	Sol profilden sağ profile çeşitli açılar.
AR [104]	1998	576×768, E	116, 3.288	Yİ: 4, AY: 4, OK: 2, KS: 2.
Georgia Tech [105]	1999	640×480, E	50, $\sim$ 1.500	Yİ, AY ile ön ve/veya eğik yüzler.
Univ. of Oulu [106]	1999	428×569, E	125, 2.000	Hepsi 16 farklı AY ile önden görüntülerdir.
CMU PIE [107]	2000	640×486, E	68, 41.368	BP: 13, AY: 43, Yİ: 3.
Human Scan [108]	2001	384×286, H	23, $\sim$ 1.500	Esas olarak önden görünümler.
Equinox Inf. [109]	2001	240×320, H	91, 25.000	Kamera spektral aralığı: 8-12 mikrometre görünürlük, AY: 3, Yİ: 3.
Yale B [110]	2001	640×480, H	10, 5.760	BP: 9, AY: 64.
Notre Dame HumanID [111]	2002	1600×1200, E	$\geq$ 300, $\geq$ 15.000	AY: 3, Yİ: 2, KS: 10-13.
Indian Face [112]	2002	640×480, H	40, 440	Parlak homojen bir arka planla ve dik, önden bir konumda deneklerle çekilmiş görüntüler.

Tablo 1.3 2003-2009 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.)

Veri seti	**Yıl**	**Resim Boyutu, Renk**	**Denek Sayısı, Resim Sayısı**	**Özellikler**
Korean Face (KFDB)[113]	2003	640×480, E	1000, 52.000	Yİ: 5, AY: 16, BP: 7.
CVL Face [114]	2003	640×480, E	114, ∼ 800	Profil sol / sağ, 45 derece sol / sağ, önden, önden gülümseme, dişlerle önden gülümseme.
CAS-PEAL [115]	2003	360×480, E	2.747, 30.900	Yİ: 6, AY: 9-15, BP: 21, AK: 6, AP: 2-4, KS:2.
FRGC(3D) [116]	2004	1704×2272 yada 1200×1600, E	466, 50.000	Kontrollü ve kontrolsüz koşullar altında çekilmiş görüntüler.
FEI Face [117]	2006	640×480, E	200, 2.800	Görüntüler, dik bir şekilde önden veya 180 dereceye kadar profil döndürerek beyaz homojen bir arka planda alınmıştır.
BU-3DFE (Static)[118]	2006	Ç, E	100, 2.500	Yİ:7, Yaş Aralığı: 18-70 , ET: 6.
MORPH [119]	2006	Ç, E	13.618, 55.134	Cinsiyet:%81 Erkek, %19 Kadın Yaş Aralığı: 16-77, ET: 4.
LFW [120]	2007	150×150, E	5.749, 13.233	Pozlanmamış fotoğraflar, çoğunlukla önden görünümler.
CASIA 3D [121]	2007	Ç, E	123, 4.624	Yİ: 6
MUCT [122]	2008	480×640, E	276, ∼ 3.500	Önden ve üç çeyrek görünümler. AY, manuel nirengi noktaları.
Bosphorus [123]	2008	Ç, E	105, 4.652	Yİ: 34, BP: 13 , OK: 4.
CMU Muti-PIE [124]	2008	Ç, E	337, ≥ 750.000	BP:15, AY:19, Bazı yüksek çözünürlüklü ön görüntüler.
CUFS [125]	2009	Ç, E	606, ∼ 1.200	Her denek için bir önden görüntü ve bir eskiz.
CASIA HFB [126]	2009	Ç, E	100, 800	Denek başına 4 VIS ve 4 NIR yüz görüntüsü.
CUFSF [127]	2009	Ç, H	1.194, ∼ 2.400	AY, Bir sanatçı tarafından çizilmiş abartılı eskiz.

Tablo 1.4 2010-2018 yılları arasında yüz tanıma için yayınlanan görüntü tabanlı veri kümeleri. (Tablo içinde kullanılan kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet), Yİ (Farklı yüz ifadelerinin sayısı), AY (Aydınlatmalar), BP (Baş pozları), OK (örtüşmeler, ör. el, saç, gözlük, sakal...), KS (Kayıt sayısı), AK (Aksesuarlar), AP (Arka planlar), ET (Etnik yapı). Derinlik, her denek için görüntü sayısının yüksek olduğu anlamına gelmektedir ve genişlik, her denek için mümkün olduğu kadar çok görüntü ile denek sayısının yüksek olduğu anlamına gelmektedir.)

Veri seti	**Yıl**	**Resim Boyutu, Renk**	**Denek Sayısı, Resim Sayısı**	**Özellikler**
CASIA NIR-VIR 2.0 [128]	2010	640×480, E	725, ∼ 36.000	Denek başına 1-22 arası VIS ve 5-50 arası NIR yüz görüntüsü.
FG-NET[129]	2011	Ç, E	82, 1.002	Yaş Aralığı:0-69, %58 Erkek, %42 Kadın.
10K US Adult Face[130]	2013	72×256, E	2.222, 10.168	Hatırlanabilirlik puanları, bilgisayar görüşü ve psikoloji özellikleri ve nirengi noktası bilgileri.
CASIA WebFace [131]	2014	Ç, E	10.575, 494.414	Ünlüler.
Cross-Age Celebrity [132]	2014	Ç, E	2.000, ≥ 160.000	16-62 yaş aralığındaki ünlüler.
CelebFaces+ [133]	2014	Ç, E	10.177, 202.599	Ünlüler, Özel.
Facebook [134]	2014	Ç, E	4.400, 4.4M	Özel veri kümesi.
Google [135]	2015	Ç, E	≥ $10M$, ≥ $500M$	Özel veri kümesi.
VGG Face [37]	2015	Ç, E	2.622, 2.6M	Ünlüler, sınırlayıcı kutular ve pozlarla ek açıklamalar
MS-Celeb-1M(Ch1) [136]	2016	Ç, E	100.000, 10M	Genişlik; Ünlüler; Veri tabanı.
MS-Celeb-1M(Ch2) [136]	2016	Ç, E	20.000, 1.5M	Genişlik; Ünlüler; Veri tabanı.
Mega Face [137]	2016	Ç, E	672.052, 4.7M	Genişlik; Bütün uzun kuyruk, Ortak nokta.
CFPW [138]	2016	Ç, E	500, 7.000	Ünlülerin ön cepheden resimleri
VGG Face2 [139]	2017	Ç, E	9.131, 3.31M	Derinlik; BP, yaş, ET; Ünlüler
MS-Celeb-1M(Ch3) [136]	2018	Ç, E	180.000, 6.8M	Denişlik; Ünlüler
CPLFW [140]	2018	250×250, E	3.968, 11.652	BP, Ünlüler

Tablo 1.5 1997-2018 yılları arasında yüz tanıma için yayınlanan video tabanlı veri kümeleri. (Kısaltmalar: Ç (Çeşitli), H (Hayır), E (Evet))

Veri Kümesi	**Yıl**	**Kare Boyutu, Renk**	**Denek Sayısı, Video Sayısı**	**Özellikler**
Univ. of Maryland [141]	1997	560×240, H	40, 70	6 farklı yüz ifadesi
XM2VTS Video [142]	1999	576×720, E	295, 1.180	Önden, profil, konuşma.
CK [143]	2000	640×480, H	97, 486	Her sekans nötr bir ifade ile başlar ve bir tepe ifadesi ile biter., Yİ: 6.
CMU MB (MoBo) [144]	2001	Ç, E	25, 100	Koşu bandının etrafında bulunan altı adet kamera kullanılarak görüntü elde edilmiştir.
Honda/UCSD [145]	2002	640×480, E	20, 1.500	Kapalı ortam, 15 sks, video uzunluğu > 15 saniye.
Texas Univ. [146]	2003	720×480, H	284, ∼ 2.500	Hareketsiz yüz çekimleri, konuşma anında dinamik yüz ifadeleri.
Max Planck [147]	2003	786×576, E	246	Altı farklı bakış açısından kaydedilen yüz hareketi birimleri.
Banca Multi-Modal [148]	2004	720×576, E	208, ∼ 2.500	Dört dil; her denek 3 aylık bir süre boyunca, 12 farklı oturumda kaydedildi.
YouTube Celeb.[149]	2008	Ç, E	47, 1910	Tüm videolar MPEG4 formatında 25 fps hızında kodlanmıştır.
BU-4DFE [150]	2008	1040×1329, E	101, 606	25 fps'de 3B veri; 6 farklı yüz ifadesi.
MOBIO [151]	2009	Ç, E	152, 1.824	Her denek için 12 seans.
YouTube Face [152]	2011	Ç, E	1.595, 3.425	YouTube'dan indirilen videolar. Her denek için ortalama 2,15 video.
PaSC Video [153]	2013	Ç, E	265, 2.802	Alternatif sensörler, önden ve önden olmayan görünümler
IJB-A [154]	2015	Ç, E	500, 2.085	Elle yerelleştirilmiş yüz görüntüleri.
IJB-B [155]	2017	Ç, E	1.845, 7.011	Elle yerelleştirilmiş yüz görüntüleri.
UMDFaces [156]	2017	Ç, E	3.107, 22.075	Genişlik; video.
IJB-C [157]	2018	Ç, E	3.531, 11.779	Bu veri kümesi 10.000 yüz olmayan görüntüye sahiptir.

(i) **Yanlış Eşleşme Oranı (FMR) (Yanlış Kabul Oranı (FAR) olarak da bilinir):** Sisteme giriş yapmaya deneyen ve sisteme erişim sağlamış olan sahtekarların, sisteme giriş yapmayı deneyen bütün sahtekarların oranına verilen isimdir.

$$\textit{Yanlış Eşleşme Oranı (FAR)} = \frac{\textit{Sisteme Kabul Edilen Sahtekar Sayısı}}{\textit{Toplam Sahtekar Örnek Sayısı}} \tag{1.1}$$

(ii) **Yanlış Eşleşmeme Oranı (FNMR) (Yanlış Reddetme Oranı (FRR) olarak da bilinir)**: Sisteme giriş yapmaya deneyen ve sisteme erişim sağlayamayan gerçek kullanıcıların, sisteme giriş yapmayı deneyen bütün gerçek kullanıcıların oranına verilen isimdir.

$$\textit{Yanlış Eşleşmeme Oranı (FRR)} = \frac{\textit{Sisteme Kabul Edilmeyen Gerçek Kullanıcı Sayısı}}{\textit{Toplam Gerçek Kullanıcı Örnek Sayısı}} \tag{1.2}$$

(iii) **Doğruluk:** Doğru sınıflandırılmış örneklerin yüzdesidir.

$$\textit{Doğruluk} = \frac{\textit{Doğru Sınıflanan Örnek Sayısı}}{\textit{Toplam Örnek Sayısı}} \tag{1.3}$$

(iv) **Gerçek Kabul Oranı (GAR) (Doğru Kabul Oranı (TAR) olarak da bilinir):** Sisteme giriş yapmaya deneyen ve sisteme erişim sağlamış olan gerçek kullanıcıların, sisteme giriş yapmayı deneyen bütün gerçek kullanıcıların oranına verilen isimdir. (yani TAR = 1-FNMR).

$$\textit{Doğru Kabul Oranı (TAR)} = \frac{\textit{Sisteme Kabul Edilen Gerçek Kullanıcı Sayısı}}{\textit{Toplam Gerçek Kullanıcı Örnek Sayısı}} \tag{1.4}$$

(v) **Tespit Oranı:** Doğru tespit edilen davetsiz misafirlerin (örneklerin değil) yüzdesidir.

(vi) **Eşit Hata Oranı (EER):** FAR ve FRR'nin eşit olduğu hata oranıdır. Biyometrik sistemlerde Eşit Hata Oranının düşük olması sistem başarımının yüksek olduğunu gösterir.

(vii) **Alıcı Çalışma Özellikleri (ROC) Eğrisi:** Bu eğri, farklı algılama eşik değerlerinde elde edilen FRR ile FAR'ın grafiğidir. ROC eğrisi, TAR'a karşı FAR'ı çizerek de elde edilebilir. ROC Eğrisinin Altındaki Alan (AUROC), 0,5

(rastgele seçim) ile 1 (mükemmel sınıflandırma) arasındaki değerleri alan, sistemin performansını temsil eden ölçüttür.

Yüz tanıma için yaygın olarak kullanılan değerlendirme ölçütleri ve çizelgeleri şunlardır: i) sıra-1 doğruluğu ve ii) k sıradaki kimlik oranının grafiği olan Kümülatif Eşleşme Karakteristik (CMC) eğrisi (doğru kimlik ilk k sonuçları arasındadır). Yukarıda bahsedilen değerlendirme mertikleri ve grafikler, yüz tanıma meydan okumalarında [160–162] ve protokollerde [157, 163] kıyaslama ve karşılaştırma amacıyla kullanılmaktadır.

1.1.5 Yüz Yanıltma Saldırıları

Yüz tanıma, kullanımı kolay bir biyometrik özellik olmasına rağmen, fotoğraflar, videolar veya 3 boyutlu maskeler kullanılarak gerçekleştirilen sahtekarlık saldırılarına karşı savunmasız olması en büyük problemlerden birini oluşturmaktadır [164–166]. Yanıltma atakları en çok biyometrik verilerin kaydedilmesi, öznitelik çıkarımı veya karar aşaması sırasında görülmektedir. Bu saldırıların yanı sıra, ağa veya biyometrik verilerin depolandığı veritabanına yapılan başka saldırı türleri de bulunmaktadır [167].

Yüz tanımada yanıltma saldırılarına karşı geliştirilen önleme çalışmaları, genellikle canlılık tespiti veya göz kırpma [168–173], yüz ifadesi değişiklikleri, ağız hareketleri [174–177] veya baş hareketleri [178] gibi fizyolojik hareketlerin algılanmasıyla yapılabilen sunum saldırısı tespiti kullanılarak gerçekleştirilmektedir. Bir yüz videosundan kalp atış hızının saptanması, canlılığın saptanması için kullanılmakta olan başka bir yöntemdir [68, 179]. Bu tekniğe uzak (temassız) fotopletismografi adı verilmektedir ve bu yöntem, kalbin her atışında ve vücuda kan pompalamasında meydana gelen derinin ince renk değişimlerini kullanmaya dayalı bir yöntemdir [180, 181].

Yüz tanıma sahtekarlığına karşı alınan diğer önlemler, yürüyüş ve konuşma gibi farklı biyometrik modaliteleri içerebilmektedir. Literatürdeki çalışmalarda incelendiği zaman, çok kipli sistemlerin [172] sahtekarlık saldırılarına karşı dayanıklılıklarının, tek kipli sistemlerle [182] kıyaslandığı zaman çok daha yüksek olduğu gözlemlenmiştir. 2 Boyutlu yüz tanıma sahtekarlığı saldırılarının üstesinden gelmek amacıyla literatürde gerçekleştirilen çalışmalar hakkında daha fazla bilgi bu konu hakkındaki derleme çalışmaları [183–188] incelenerek elde edilebilir.

DESA tabanlı yöntemler son zamanlarda, literatürde bulunan birçok problemin çözümünde olduğu gibi yüz sahtekarlığına karşıda popüler hale gelmiştir [166, 189].

Nagpal ve arkadaşları tarafından gerçekleştirilen çalışmada [190], yüz sahtekarlığı önleme için kullanılan farklı evrişimli sinir ağları mimarilerinin performansları incelenmiştir. Liu ve arkadaşları tarafıdan [191], sıfır vuruşlu yüz yanıltma saldırıları önleme (ZSFA) gerçekleştirebilmek için bir derin ağaç öğrenme (DTL) yöntemi önerilmiştir. ZSFA, kısmi kağıt veya şeffaf maske saldırıları gibi eğitim verilerinde bulunmayan yanıltma saldırılarının tespiti için önerilen metotlara verilen genel addır. Yang ve arkadaşları tarafından 2019 yılında [192], sahte yüzleri tespit etmek için sınırlar ve hareli desenler gibi ince ipuçlarına odaklanabilen bir uzay-zaman sahtekarlığı önleme ağı (STASN) önerilmiştir. Ayrıca, yine bu çalışmada, yüz yanıltma saldırıları önleme ile ilgili veri toplama ve sentez çözümleri sunmuşlardır.

Son zamanlarda CASIA-SURF çok kipli veri kümesi [193] kullanılarak çok kipli bir yüz yanıltma saldırıları önleme meydan okumasıda gerçekleştirilmiştir. Bu çalışmada, en başarılı ekiplerin sonuçlarını özetlemiş ve gelecekteki araştırma yönleri hakkında önemli bilgiler verilmiştir [194].

1.1.6 Derin Öğrenme Temelli Yüz Tanıma Sistemleri

Derin sinir ağları (DSA), işlem gücündeki artış ve (çoklu) etiketli örnekleri içeren büyük veri kümelerinin derlenebilmesi sayesinde 2014 yılından beri yüz tanıma sistemlerinin oluşturulmasında başarıyla kullanılmaktadır [36, 135, 195–197]. DESA'ya dayalı DeepFace yöntemi [36], sınırsız koşullar altında çekilen binlerce yüz görüntüsünü içeren Labeled Faces in the Wild (LFW) veri kümesi [120] kullanılarak gerçekleştirilen deneysel çalışmalar sonucunda yaklaşık %97,35'lik bir yüz tanıma başarımı sağlamıştır, ve bu elde edilen yüz tanıma başarımı insanın yüz tanıma seviyesine çok yakın bir değer olarak kayıtlara geçmiştir (%97.53). O yıldan itibaren gerçekleştirilen çalışmalar sonucunda, LFW veri kümesi kullanılarak elde edilen yüz tanıma başarımları %99.80'lere ulaşmıştır. Derin öğrenme tabanlı yöntemler doğrusal olmayan görünüm tabanlı yöntemler olmasına rağmen, yüz tanıma çalışmalarının çoğunda derin öğrenme tabanlı yöntemlerin kullanılmış olması nedeniyle, bu tez çalışmasında derin öğrenme tabanlı yöntemler yeni bir alt başlık altında incelenmektedir.

Derin öğrenme tabanlı yüz tanıma yöntemleri temel olarak üç aşamadan oluşur [19]: Yüz ön işleme, derin öznitelik çıkarma ve yüz eşleştirme. Aşağıdaki bölümlerde bu adımların her biri hakkında kısa bilgiler verilmektedir:

Yüz Ön İşleme: Derin öğrenme tabanlı yüz tanıma sistemleri, kontrolsüz ortamlarda farklı aydınlatma, poz ve yüz ifadeleriyle yüz görüntülerini tanımada belirli bir sağlamlık sağlamaktadır. Ancak son dönemlerde gerçekleştirilen çalışmalar [14, 198],

farklı aydınlatma, pozlama ve yüz ifadelerinin ağın yüz tanıma performansı üzerinde olumsuz etkileri olduğunu göstermiştir ve bu durumlar altında ağın yüz tanıma performansını artırmak için yüz ön işlemeye ihtiyaç olduğu ortaya çıkmıştır.

Birden çok artırma (one-to-many augmentation) ön işleme yaklaşımında amaç, eğitim sırasında tek bir görüntüden farklı pozlarda görüntüler üreterek poz değişiminden etkilenmeyen DESA'lar oluşturmaktır. Bu ön işleme sürecinin gerçekleştirilmesinin amacı, bir eğitim veri tabanı oluşturmak için çok sayıda görüntü toplamanın pahalı ve zaman alıcı bir işlem olmasıdır. Birden çok artırma ön işleme için kullanılan yaklaşımlardan ilki, fotometrik dönüşümler [35] ve geometrik dönüşümler [199, 200] gibi veri artırma yöntemlerini kullanılmasıdır. İkinci yaklaşım, farklı pozlarda 2 boyutlu görüntüler kullanılarak bir 3 boyutlu yüz modelinin yeniden oluşturulmasına dayanır [201, 202]. Üçüncü yaklaşımda, DESA modelleri, 2 boyutlu görüntülerden 3 boyutlu modeller oluşturmak ve bunları farklı pozlardaki 2 boyutlu görüntülere geri yansıtmak yerine, doğrudan 2 boyutlu görüntüler oluşturmak için kullanılır [203, 204]. Bao ve arkadaşları [205], bir giriş görüntüsünden ve diğer herhangi bir giriş görüntüsünden elde edilen özniteliklerden yeni bir yüz sentezlemişler ve bu sentezin eğitim veri kümesi içinde olmasa bile gerçekçi yüz görüntüleri üretmede çok başarılı olduğunu göstermişlerdir. Üretken Çekişmeli Ağları (ÜÇA)'da bu amaçla kullanılmaktadır [206–209].

Çoktan bire normalleştirme (many-to-one normalization) ön işleme yaklaşımında amaç, kontrolsüz ortamlarda farklı açılardan elde edilen yüz görüntülerini kullanarak yüz görüntüsünün kanonik görüntüsünü oluşturmaktır. Yığılmış Otomatik Kodlayıcılar (YOK) [200], ESA [210] ve ÜÇA [211] yapıları, farklı açılara sahip çoklu görüntülerden yamalar kullanarak bir ön yüz görüntüsü elde etmek için kullanılmıştır.

Son zamanlarda gerçekleştirilen bir çalışmada [212], görünüm-akış tabanlı ESA'ya dayalı bir yüz cepheleştirme yöntemi önerilmiştir. An ve arkadaşları tarafından gerçekleştirilen çalışmada ise [213], yüz pozlarını kullanarak hizalama şablonlarını uyarlanabilir bir şekilde öğrenen ve ardından bu şablonları kullanarak test ve eğitim görüntülerini hizalayan bir uyarlamalı poz hizalama yöntemi önerilmiştir. Son dönemde gerçekleştirilen bir çalışmada ise, yumuşak ve sert gölgeleri ortadan kaldıran ve kimlikle ilgili bilgileri tutan, aydınlatma değişimlerine dayanıklı bir ön işleme yöntemi önerilmiştir [214].

Derin Öznitelik Çıkartma: Öznitelik çıkartma işleminde kullanılacak olan DESA'ların tasarlanması aşamasında karar verilmesi gereken en önemli iki parametre ağ mimarisinin ve kayıp fonksiyonunun belirlenmesidir. Derin evrişimli sinir ağlarının mimari yapıları, omurga mimarileri ve çoklu mimariler olmak üzere iki ana başlık

altında toplanabilir. Tipik CNN mimarileri olarak bilinen AlexNet [35], VGGNet [215], GoogleNet [216], ResNet [217] ve SENet [218] gibi derin öğrenme ağları ImageNet [219] yarışmalarındaki yüksek performansları sonrasında araştırmacıların ilgisini çekmiştir. Daha sonraki süreçlerde, bu ağlar ve varyasyonları, nesne tanımanın yanında yüz tanıma içinde kullanılmıştır. Ana akım derin sinir ağlarına ek olarak, yüz tanıma dahil olmak üzere çoklu göreve sahip öğrenme gerçekleştirmek amacıyla çoklu mimariye sahip ağlar önerilmiştir [94, 220].

Kayıp fonksiyonunun seçimi, DESA'ların yüz tanıma için eğitilmesi için mimarilerin seçimi kadar önemli, karar verilmesi gereken diğer bir parametredir. Yüz tanıma işlemlerinde, sınıf içi farklılıklar sınıflar arası farklılıklardan daha büyük olduğu için, öznitelikleri ayırmak için softmax kayıp fonksiyonunun yeterli olmadığı görülmüştür. Bu nedenle, öznitelikleri daha ayırt edici hale getirmek için Öklid mesafesine dayalı kayıp [221, 222], üçlü kayıp [135], açısal/kosinüs marjı tabanlı kayıp [223] ve yumuşak-maksimum kayıp [224, 225] gibi farklı kayıp fonksiyonları önerilmiş ve önerilen bu kayıp fonksiyonları yüz tanıma sistemleri için kullanılmıştır.

Öznitelik çıkarma, özellikle tek çekim yüz tanıma probleminin çözümü için hala geliştirilmesi gereken bir işlemdir. Tek çekim (veya düşük çekim) yüz tanıma, galerideki bazı öznelerin yalnızca tek bir (veya birkaç) görüntüsünün olduğu [226–230], varyanslarının birkaç örnekle temsil etmenin zor olduğu durumlar nedeniyle hala açık bir araştırma problemi olarak durmaktadır. Literatürde tek çekim yüz tanıma problemini çözmek amacıyla, çeşitli yaklaşımlar önerilmiştir. Jadhav ve arkadaşları tarafından 2016 yılında önerilen bir yöntemde [228], tamamen yüz tabanlı öznitelik temsillerinden daha iyi performans gösterdiği tespit edilen cinsiyet ve yüz şekli gibi belirli öznitelikler için bir DESA'nın ince ayarıyla elde edilmiş ara derin öznitelik temsilleri kullanılmıştır. Guo ve Zhang tarafından bu problemin çözümü ile ilgili [227], çapraz entropi kayıp fonksiyonu ile birlikte bir düzenlileştirme fonksiyonu kullanılmış ve yeni bir az temsil edilen sınıflar promosyon kayıp fonksiyonu önerilmiştir. En iyi sonuçlar, Bilgisayarlı Görme Uluslararası Konferans (ICCV) 2017'deki MS-Celeb-1M Düşük Çekim Yüz Tanıma Yarışması'nda elde edilmiştir. Yeni bir düzenlileştirme terimi kullanan başka bir yöntem Wu ve arkadaşları tarafından önerilmiştir [230]. Sınıflandırıcının tek seferlik sınıfları daha iyi öğrenmesini sağlamak amacıyla Ding ve arkadaşları tarafından 2018 yılında gerçekleştirilen çalışmada [226], tek seferlik sınıflar için anlamlı verileri sentezlemek için üretken bir çekişmeli ağ (GAN) kullanıldığı farklı bir yaklaşım önerilmiştir.Yeni örnekler, daha fazla örneğe ait diğer sınıflardan gelen veri varyansları kullanılarak uyarlanması ile oluşturulmuştur.

Yüz eşleştirme: Kararlaştırılan kayıp fonksiyonu ve eğitim verileri kullanılarak, mimari yapısına karar verilmiş olan derin öğrenme ağı eğitildikten sonra, test verileri için derin öznitelik çıkarımı yapılabilmekte, yüz tanıma ve doğrulama işlemleri gerçekleştirilebilmektedir. Yüz eşleştirme, kosinüs mesafesi veya L_2 mesafesi kullanılarak gerçekleştirilebilmektedir. Metrik öğrenme ve Seyrek Temsil Tabanlı Sınıflandırıcı (STS) [19] gibi yöntemler de yüz eşleştirme için kullanılan diğer metotlara örnektir.

Taigman ve arkadaşları tarafından gerçekleştirilen çalışma [36], nesne tanımada başarılı olan derin öğrenme ağlarının yüz tanımada da çok başarılı olacağını göstermiştir. DeepFace olarak bilinen bu çalışmada, softmax kayıp fonksiyonu ile ağ mimarisi olarak Alexnet kullanılmıştır. Önerilen yöntemin, Facebook veri kümesindeki test performansı, insan yüz tanıma performansına yakın olan %97.35 idi. 2014 yılında yapılan bir başka çalışma [133], AlexNet'in ağ mimarisi olarak kullanıldığı ve karşılaştırmalı kayıp fonksiyonun kullanıldığı metot "DeepID2" olarak adlandırıldı. CelebFaces+ veri kümesi kullanılarak gerçekleştirilen testlerdeki doğruluk oranı %99,15 idi. Bir yıl sonra, ağ mimarisi olarak AlexNet mimarisi yerine VGG-Net10'un ağ mimarisi olarak kullanıldığı metota "DeepID3" adı verildi ve %99,53'lük bir test performansı elde edilmiştir [231].

Schroff ve arkadaşları [135], GoogleNet-24'ü ağ mimarisi olarak kullanıldığı üçlü kayıp fonksiyonun kayıp fonksiyonu olarak belirlendiği ve ağın eğitimi için Google veritabanı kullandıkları FaceNet'i 2015 yılında yüz tanıma için önermişlerdir. Google veritabanı, o zamana kadar yüz tanıma için kullanılan en büyük veri kümesiydi. Üçlü kayıp fonksiyonu, bir bağlantı görüntüsünü, aynı sınıfın olumlu bir örneğini ve farklı bir sınıfın olumsuz bir örneğini içermektedir. Amaç, ESA'yı, eşleşen çift arasındaki mesafeyi en aza indirecek ve eşleşmeyen çift arasındaki mesafeyi en üst düzeye çıkaracak şekilde güncellemektir. FaceNet'in LFW veri kümesi kullanılarak gerçekleştirilen test çalışmalarında doğrulama performansı %99,63 olarak elde edilmiştir.

Parkhi ve arkadaşları [37], VGGNet-16'yı ağ mimarisi olarak kullanıldığı üçlü kayıp fonksiyonun kayıp fonksiyonu olarak belirlendiği ve ağın eğitimi için VGGFace veri kümesi kullandıkları VGGFace'i 2015 yılında yüz tanıma için önermişlerdir. LFW veri tabanı kullanılarak gerçekleştirilen testlerde, %98,95 doğrulama performansı elde edilmiştir. 2016'da yüz tanıma için önerilen yöntemlerden biri, kayıp fonksiyonu olarak merkez kayıp fonksiyonunu kullanmış ve LFW veri kümesi kullanılarak gerçekleştirilen testlerde daha yüksek performans elde etmiştir [232]. 2017 yılında önerilen yöntemlerin çoğu ResNet mimarisini ve onun varyasyonlarını [233–238] tercih etmiştir. Liu ve diğerleri tarafından önerilen yöntemde [235], eğitim için

MS-Celeb-1M veri kümesini [136] ve CoCo (türdeş kosinüs) kayıp fonksiyonunu kullanmıştır.LFW veri kümesi kullanılarak gerçekleştirilen test çalışmalarında elde edilen %99,86'lık doğrulama başarımı, o zamana kadar elde edilen en yüksek doğrulama başarımıdır. LFW veri tabanı kullanılarak gerçekleştirilen yüz tanıma test performanslarına bakıldığında, son 5 yılda yüz tanıma sistemlerinin performansında olağanüstü bir artış olduğu (%97,35'ten %99,86'ya) gözlemlenmektedir.

1.2 Tezin Amacı

Yüz tanıma sistemlerinin geliştirilmesi, 1980'li yıllardan itibaren literatürde araştırmacıların ilgisini çeken bir problem olmuştur. Özellikle 2010'lu yılların başına kadar, bu problemin çözümü ile ilgili olarak birçok yöntem önerilmiş ve önerilen yüz tanıma sistemlerinin doğrulama ve tanıma başarımları kontrollü ortamlarda %90'lara ulaşmıştır. 2010'lu yıllarında başında derin öğrenme ağlarının nesne tanımada çok büyük sayıda veriye ve sınıfa sahip veri kümesi kullanılarak gerçekleştirilen deneysel çalışmalarda yüksek tanıma başarımı göstermesi ile birlikte, literatürde yüz tanıma gibi birçok problemin çözümünde derin öğrenme ağlarının kullanımı yaygınlaşmıştır. Taigman ve arkadaşları tarafından gerçekleştirilen çalışma [36], nesne tanımada başarılı olan derin öğrenme ağlarının yüz tanımada da çok başarılı olacağını göstermiştir. Taigman tarafından önerilen yöntemin, Facebook veri kümesindeki test performansı, insan yüz tanıma performansına yakındır. Bu elde edilen yüz tanıma performansı ile, yüz tanıma problemi bir meydan okuma haline gelmiştir ve literatürde yüz tanıma için birçok yeni yöntem ve algoritma önerilmiştir. LFW veri tabanı kullanılarak gerçekleştirilen yüz tanıma test performanslarına bakıldığında, son 5 yılda yüz tanıma sistemlerinin performansında olağanüstü bir artış olduğu (%97,35'ten %99,86'ya) gözlemlenmektedir. Elde edilen test sonuçları, artık yüz tanıma işleminin artık yüksek başarımla gerçekleştirilebilen bir bilgisayarlı görme problemi olduğunu ortaya koymaktadır.

Tüm elde edilen bu yüz tanıma performansları sonucunda, araştırmacılar farklı koşullar altında derin öğrenme ağlarının yüz tanıma performanslarını incelemişlerdir. Derin öğrenme temelli yüz tanıma yöntemleri, kontrollü ortamlar altında hem yüz tanımlama hem de doğrulama görevleri için yüksek performans elde etmelerine rağmen, performanslarının olumsuz koşullarda (örneğin aydınlatma, kontrast ve gürültü varyasyonları) [14, 15] önemli ölçüde azaldığı görülmüştür. Olumsuz koşullar litaratürde ayrıca semantik saldırılar olarakta adlandırılmaktadır [16]. Bu tezde, olumsuz koşullar altında yüz tanıma başarımı azalan derin öğrenmeye dayalı yüz tanıma sistemlerinin, yüz tanıma başarımını artırmak amacı ile yumuşak biyometrilerin yüz tanıma sistemine yardımcı veriler olarak kullanılması ve bu

kullanılan yumuşak biyometri verileri sayesinde yüz tanıma sisteminin başarımının artırılması amaçlanmıştır.

1.3 Hipotez

Bu tezin amacı, olumsuz koşullar altında yüz tanıma başarımı azalan derin öğrenmeye dayalı yüz tanıma sistemlerinin, yüz tanıma başarımını artırmak amacı ile yumuşak biyometrilerin yüz tanıma sistemine yardımcı veriler olarak kullanılması ve bu kullanılan yumuşak biyometri verileri sayesinde yüz tanıma sisteminin başarımının artırılmasıdır. Bu amaç doğrultusunda, öncelikle yüze ait yumuşak biyometrilerin nasıl elde edileceği ile ilgili literatür araştırması gerçekleştirilmiş ve yüze ait dinamik özniteliklerin elde edilmesi ile ilgili bir yöntemin kullanılmasına karar verilmiştir.

Daha sonra karar verilen yöntem ile elde edilen yüze ait dinamiklerin kişinin kimliğine ait bilgi taşıyıp taşımadığı incelenmiştir. Yüze ait dinamiklerden kişinin kimlik bilgilerine dair bilgi taşıdığı gösterilmiş, derin öğrenme temelli yüz tanıma sistemlerinin hangi zorlu şartlar altında yüz tanıma başarımının düştüğü incelenmiştir. Daha sonra farklı koşullar altında hem yüze ait dinamiklerden elde edilen öznitelikler hem de derin öğrenme ağlarından elde edilen öznitelik vektörleri birlikte kullanılarak, derin öğrenme temelli yüz tanıma sistemlerinin başarımının düştüğü zorlu koşullar altında, önerilen yeni yüz tanıma sisteminin başarımının derin öğrenme ağları temelli yüz tanıma sistemlerine kıyasla daha yüksek olduğu gösterilmiştir.

2
YÜZ HAREKETLERİNDEN YUMUŞAK BİYOMETRİK VERİ ÇIKARILMASI

Literatürdeki yüz tanıma yöntemleri, genel olarak görüntü tabanlı ve video tabanlı yöntemler olarak ikiye ayrılmaktadır. Görüntü tabanlı yöntemler genellikle yüz görünüm özelliklerini kullanırken, video tabanlı yöntemlerde yüzün görünüm ve dinamik özellikleri birlikte kullanılabilmektedir. İyi kalitede görüntüler kullanılarak, kontrollü ortamlarda yüksek yüz tanıma oranları elde edilebilmesine rağmen, yüz ifadesi, baş pozisyonu, aydınlatma değişiklikleri, örtüşme, düşük çözünürlük ve gürültü gibi daha gerçekçi ve zorlu koşullar altında en iyi performans gösteren derin öğrenme tabanlı yöntemlerde dahi yüz tanıma oranları daha düşüktür. Bu zorlukların üstesinden gelebilmek için yumuşak biyometrik özelliklerin kullanılması önem kazanmıştır [1]. Yumuşak biyometrik özellikler, bir kişinin anatomik veya davranışsal özelliklerini kullanarak yaş, cinsiyet, etnik köken, saç rengi, göz rengi, yürüyüş ve yüz hareketi gibi kimlik hakkında bazı bilgiler sağlamaktadır. Yüz hareketinden elde edilen yumuşak biyometrik özellikler, yaş tahmini [2], cinsiyet tahmini [3] ve yüz tanıma [50, 239, 240] için kullanılmıştır. Yüz dinamikleri, gözlük, sakal, makyaj gibi görünüm değişikliklerinden ve aydınlatmadan daha az etkilenmektedir.

Yüz tanıma için yüz dinamiklerini kullanan ilk çalışmalardan birinde, bireyler film izlerken veya röportaj yaparken tespit edilen bazı yüz hareket birimlerinin sıklığı analiz edilmiştir. Bu sayede, eylem birimi tabanlı dinamik özniteliklerin yüz tanıma için kullanılabileceği gösterilmiştir. Liu ve Cheng tarafından gerçekleştirilen bir çalışmada [241], yüzün zamansal dinamikleri ile uyarlanabilir gizli Markov modelleri kullanılarak kişi tanıma gerçekleştirilmiştir. Matta ve Dugelay tarafından gerçekleştirilen çalışmalarda ise [242, 243], katı kafa yer değiştirmelerinden elde edilen davranışsal bilgilere ve klasik Eigenface yaklaşımına [22] benzer fizyolojik bilgilere dayanan bir kişi tanıma sistemi önerilmiştir. Sınıflandırıcı yapısı olarak, Gauss Karışım Modellerini (GMM) ve Bayes sınıflandırıcısının kullanıldığı çalışmada, davranışsal ve fiziksel özellikleri birleştirmenin yüz tanıma oranını arttırdığı

gösterilmiştir. Zafeiriou ve Pantic [244] tarafından gerçekleştirilen çalışmada, doğal gülümseme sırasında yüzde meydana gelen değişikliğin biyometrik bilgi taşıdığı gösterilmiştir. Yüz hareketi, gülümseme yüz ifadesi videosunun nötr çerçevesi ve tepe çerçevesi arasında yoğun bir hareket alanı olarak temsil edilmiştir. 22 farklı kişiden elde edilen 563 gülme videosundan oluşan veri tabanı kullanılarak yapılan deneysel çalışmalarda elde edilen sonuçlar umut vericidir.

Literatürde yüz dinamiklerinden yararlanan video tabanlı yüz tanıma yöntemleri incelendiği zaman, dört farklı yaklaşım kullanıldığı gözlemlenmiştir:

(i) Yüz hareket birimlerini kullanan yöntemler [240, 245–247]

(ii) Yüzün optik akışı veya yoğun hareket alanı kullanan yöntemler [244, 248, 249]

(iii) Yüz nirengi noktalarını kullanan yöntemler [250, 251]

(iv) Yüz hareketini örtük olarak modelleyen uzamsal-zamansal yöntemler [50, 239, 252, 253]

Bu yöntemlerden sadece bazıları duygusal yüz ifadelerinin dinamiklerine odaklanmakta ve sınırlı sayıda denek (en fazla 95 denek) içeren veri kümeleri üzerinde sonuç bildirmektedir.

Bu bölümde, gülümseme videolarının duygusal yüz ifadelerinden elde edilen dinamik öznitelikler kullanılarak yüz tanıma gerçekleştirilen bir yöntem öneriyoruz. Daha önce yaş ve cinsiyet tahmini için [2, 3] önerilen bir nirengi noktası tabanlı dinamik öznitelik çıkarma yöntemi benimsenmiştir.

Bu bölümde, öncelikle önerilen yöntem hakkında detaylı bilgi verilecek, UvA-NEMO gülümseme veri kümesi kullanılarak gerçekleştirilen deneysel çalışmaların sonuçları verilecektir. Sonrasında elde edilen sonuçlar ışığında, yüze ait dinamiklerden elde edilen özniteliklerin kişiye ait kimlik bilgisi taşıyıp taşımadığı tartışılacaktır.

2.1 Yüze Ait Dinamik Öznitelikler Kullanılarak Kimlik Bilgisi Doğrulama

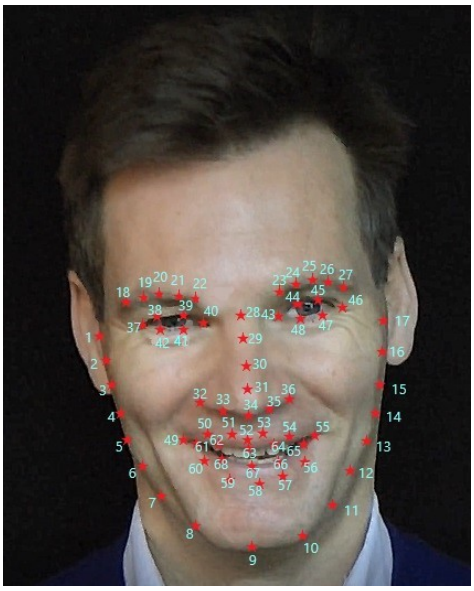
Yüze ait dinamikleri kullanarak yüz tanıma gerçekleştirmek için önerilen yöntem, Şekil 2.1'de gösterildiği gibi dört adımdan oluşmaktadır. Bu adımlar; yüz görüntüsünün elde edilmesi, yüz nirengi noktalarının tespit edilmesi, yüz nirengi noktalarını kullanarak dinamik özniteliklerin çıkarılması ve benzerlik puanı olarak Öklid mesafelerini kullanarak yüz doğrulamasının gerçekleştirilmesidir.

Şekil 2.1 Yüz dinamiklerine dayalı yüz tanıma yönteminin adımları.

2.1.1 Yüz Nirengi Noktalarının Tespit Edilmesi

Videodan yüz tanıma algoritmalarının ilk adımı, video karelerindeki yüzü algılamak ve izlemektir. Yüz bölgesinin tespiti, başın pozisyonuna, oklüzyona ve aydınlatma değişikliklerine dayanıklı olacak şekilde yapılmalıdır. Çok çeşitli baş pozisyonları için yüz ve yüz işaretlerini tespit edebilen yakın zamanda önerilen birkaç algoritma vardır [79, 80, 88, 254, 255]. Ağaç yapısı modeline dayalı yüz ve nirengi noktası algılama algoritması Zhu ve Ramanan [255] tarafından önerilmiştir. Yüzlerce yüz görüntüsü üzerinde eğitilen algoritmanın, milyarlarca yüz görüntüsü ile eğitilmiş ticari bir yazılımla karşılaştırılabilir sonuçlar verdiği gösterilmiştir. CHEHRA [254] olarak bilinen başka bir çalışmada, dönüm noktalarını yüksek hassasiyetle belirlemek için her kullanıcıya özel olacak şekilde genel eğitimli bir modelin benimsenmesi önerilmiştir. IntraFace'de [80], doğrusal olmayan en küçük kareler fonksiyonunun optimizasyonu sırasında denetimli bir iniş yöntemi kullanılır ve önerilen bu yöntem yüz hizalama problemine uyarlanmaktadır.

Bu tez çalışmasında, yüz nirengi noktalarını tespit etmek amacıyla Kazemi ve arkadaşları tarafından önerilen yöntem kullanılmaktadır [79]. Kazemi ve arkadaşları tarafından önerilen yöntem, kare hata kaybının toplamını optimize eden ve doğal olarak eksik verileri yöneten bir regresyon ağaçları topluluğunu öğrenmek için gradyan artırmaya dayalı bir yöntemdir. UvA-NEMO veri tabanından elde edilmiş ve tespit edilen 68 nirengi noktasını içeren örnek bir video karesi, nirengi noktaları numaralarıyla birlikte Şekil 2.2'te gösterilmiştir. Bu nirengi noktalarının konumu kullanılarak dinamik öznitelikler çıkarılacaktır.



Şekil 2.2 Kazemi ve arkadaşları tarafından önerilen yöntem kullanılarak tespit edilen 68 nirengi noktası ile UvA-NEMO veri tabanından bir örnek video karesi.

2.1.2 Yüze Ait Dinamik Özniteliklerin Çıkarılması.

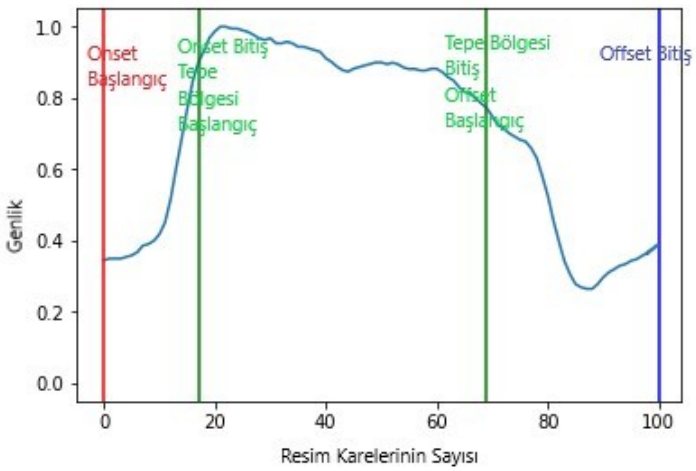
Kazemi ve arkadaşları tarafından yöntemin kullanılması ile yüze ait nirengi noktalarının konumlarının belirlenmesinin ardından, yüze ait dinamik özniteliklerin çıkartılması amacı ile ilk olarak, gülümseme eylemi sırasında hareket etmesi beklenen nirengi noktaları kullanılarak 27 adet yüz mesafesi hesaplanır. Hesaplanan 27 adet yüz mesafesi Tablo 2.1'de listelenmiştir, burada ρ iki yüz nirengi noktası arasındaki Öklid mesafesini, l_i ise i'nci yüz nirengi noktasını belirtmektedir. Her bir video karesi için Tablo 2.1'de verilen mesafeler hesaplandıktan sonra, her bir mesafe için istatistiksel özellikler hesaplanacaktır.

İnsan gülüşü, dudak köşelerinin yanındaki elmacık majör kasının kasılması ile gerçekleşmektedir [256] ve bu gülümseme aktivitesi üç ana bölümden oluşur; başlangıç, tepe bölgesi ve bitiş. Bu çalışmada, gülümseme aktivitesinin üç farklı parçaya bölünmesi, duygunun oluştuğu videodaki kare aralıklarını belirlemek amacıyla kullanılmıştır. Başlangıç bölgesi, yüz ifadesinin nötr durumdan ifade durumuna geçtiği videonun bölümünü ifade etmektedir. Tepe noktası bölgesi, ifade durumunun en yüksek dönemini gösterir ve bitiş bölgesi, yüzün ifade durumundan nötr duruma geçtiği dönemi ifade etmektedir. UvA-NEMO veri kümesindeki her video, nötr bir durumla başlar, gülümsemenin zirvesine ulaşır ve nötr durumla sona ermektedir. Başlangıç-tepe bölgesi-bitiş bölütlemesi, normalleştirilmiş ağız uzunluğu D_5 kullanılarak gerçekleştirilmiştir. Video karesi numarasına karşı normalleştirilmiş ağız uzunluğunun örnek bir grafiği Şekil 2.3'te gösterilmektedir.

Tahmini video bölümleri Şekil 2.3'te belirtilmiştir. Videonun ilk karesinden başlayarak

Tablo 2.1 Dinamik öznitelikleri çıkarmak için kullanılan yüz mesafeleri. (ρ iki yüz nirengi noktası arasındaki Öklid mesafesini, l_i ise i'nci yüz nirengi noktasını belirtmektedir.)

Yüze Ait Uzaklık	**Tanımlama**	**Yüz Nirengi Noktaları ile Tanımlama**
D_1	Sağ gözün genişliği	$\rho(\frac{l_{44}+l_{45}}{2}, \frac{l_{47}+l_{48}}{2})$
D_2	Sol gözün genişliği	$\rho(\frac{l_{38}+l_{39}}{2}, \frac{l_{41}+l_{42}}{2})$
D_3	Sağ gözün uzunluğu	$\rho(l_{43}, l_{46})$
D_4	Sol gözün uzunluğu	$\rho(l_{37}, l_{40})$
D_5	Ağız uzunluğu	$\rho(l_{49}, l_{55})$
D_6	Ağız genişliği	$\rho(l_{52}, l_{58})$
D_7	Ağız merkezinin üst dudağın sol tarafına uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, l_{51})$
D_8	Ağız merkezinin üst dudağın sağ tarafına uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, l_{53})$
D_9	Ağız merkezinin sol dudak köşesine uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, l_{49})$
D_{10}	Ağız merkezinin sağ dudak köşesine uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, l_{55})$
D_{11}	Ağız merkezinin üst dudağa uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, l_{63})$
D_{12}	Ağız merkezinin iki dudak köşesinin ortasına uzaklığı	$\rho(\frac{l_{52}+l_{58}}{2}, \frac{l_{49}+l_{55}}{2})$
D_{13}	Sağ kaşın sol tarafının buruna uzaklığı	$\rho(l_{23}, l_{31})$
D_{14}	Sol kaşın sağ tarafının buruna uzaklığı	$\rho(l_{22}, l_{31})$
D_{15}	Sağ kaşın ortasının sağ kaşın sağ tarafına uzaklığı	$\rho(l_{25}, l_{27})$
D_{16}	Sağ kaşın ortasının sağ kaşın sol tarafına uzaklığı	$\rho(l_{23}, l_{25})$
D_{17}	Sol kaşın ortasının sol kaşın sağ tarafına uzaklığı	$\rho(l_{20}, l_{22})$
D_{18}	Sol kaşın ortasının sol kaşın sol tarafına uzaklığı	$\rho(l_{18}, l_{20})$
D_{19}	Kaşlar arası uzaklık	$\rho(l_{22}, l_{23})$
D_{20}	Sol gözün sol köşesinin sol ağız köşesine uzaklığı	$\rho(l_{37}, l_{49})$
D_{21}	Sol gözün sağ köşesinin ağzın ortasına uzaklığı	$\rho(l_{40}, \frac{l_{63}+l_{67}}{2})$
D_{22}	Sağ gözün sol köşesinin ağız merkezine uzaklığı	$\rho(l_{43}, \frac{l_{63}+l_{67}}{2})$
D_{23}	Sağ gözün sağ köşesinin sağ ağız köşesine uzaklığı	$\rho(l_{46}, l_{55})$
D_{24}	Sol gözün üst kısmının sol kaşın sağ köşesine uzaklığı	$\rho(\frac{l_{38}+l_{39}}{2}, l_{22})$
D_{25}	Sağ gözün üst kısmının sağ kaşın sol köşesine uzaklığı	$\rho(\frac{l_{44}+l_{45}}{2}, l_{23})$
D_{26}	Sol gözün üst kısmının sol kaşın sol köşesine uzaklığı	$\rho(\frac{l_{38}+l_{39}}{2}, l_{18})$
D_{27}	Sağ gözün üst kısmının sağ kaşın sağ köşesine uzaklığı	$\rho(\frac{l_{44}+l_{45}}{2}, l_{27})$



Şekil 2.3 Gülüşün bölümlere ayrılmış başlangıç, tepe noktası ve bitiş aşamaları. Çizim, videodaki kare numarasına karşı normalleştirilmiş ağız uzunluğunu göstermektedir.

ağzın uzunluğunu ve iki kare arasındaki artışını izliyoruz. Ağız uzunluğu belirli bir eşiğin (0.8) altında olduğu ve sabit bir artış (%ϵ) olduğu sürece, o bölge başlangıç bölgesi olarak işaretlenmektedir. Ağız uzunluğu, eşiğin üzerindeyse ancak sürekli bir değişiklik yoksa o bölge tepe bölgesi olarak işaretlenmektedir. Son olarak, ağız uzunluğu belirlenen eşiğin altındaysa ve sürekli bir azalma varsa, o bölge bitiş bölgesi olarak belirlenmektedir. Gülümseme ifadesi başlangıç, tepe noktası ve bitiş bölütlerine ayrıldıktan sonra, Tablo 2.2'de açıklandığı üzere 27 yüz mesafesinin her biri için 24 istatistiksel dinamik öznitelik hesaplanmıştır. Üst simgeler ($^{+}$), (^{a}) ve ($^{-}$) gülümseme ifadesinin sırasıyla başlangıç, tepe noktası ve bitiş bölümlerini temsil etmektedir. Tabloda bulunan hız $V = \frac{dD}{dt}$ olarak hesaplanmakta ve ivme $A = \frac{d^2D}{dt^2}$ olarak hesaplanmaktadır. D uzaklığa bağlı öznitelikleri, η parametresi kullanılan çerçeve sayısını temsil etmekteyken, ω videonun kayıt edilmesi sırasındaki çerçeve hızını temsil etmektedir.

27 yüz mesafesinin her biri için istatistiksel dinamik öznitelikler hesaplandıktan sonra, UvA-NEMO gülümseme veri kümesindeki her bir video için $27 \times 24 = 648$ boyutlu bir öznitelik vektörü elde edilmektedir.

2.1.3 UvA-Nemo Veri Kümesi Üzerinde Yapılan Yüz Doğrulama Deneyleri

Dinamik özniteliklerin çıkartılması işlemi tamamlandıktan sonra, UvA-NEMO veri tabanı kullanılarak deneysel çalışmalar gerçekleştirilerek, elde edilen öznitelik vektörlerinin kişinin kimliği ile ilgili bilgi taşıyıp taşımadığı kontrol edilmiştir.

Tablo 2.2 İstatistiksel Yüz Dinamiği Öznitelikleri.

Öznitelik	**Tanımlama**			
	Genel	**Başlangıç**	**Tepe Noktası**	**Bitiş**
Süre		$\eta(\frac{D^+}{\omega})$	$\eta(\frac{D^a}{\omega})$	$\eta(\frac{D^-}{\omega})$
Süre Oranı		$\frac{\eta(D^+)}{\eta(D)}$		$\frac{\eta(D^-)}{\eta(D)}$
Maksimum Genlik	Maks.(D)			
Genlik Standart Sapma (SS)	SS(D)			
Ortalama Genlik		Ort.(D^+)	Ort.(D^a)	Ort.(D^-)
Toplam Genlik		$\sum(D^+)$		$\sum(D^-)$
Net Genlik	$\sum(D^+)$-$\sum(D^-)$			
Genlik Oranı		$\frac{\sum(D^+)}{\sum(D^+)+\sum(D^-)}$		$\frac{\sum(D^-)}{\sum(D^+)+\sum(D^-)}$
Maksimum Hız		Maks.(V^+)		Maks.(V^-)
Ortalama Hız		Ort.(V^+)		Ort.(V^-)
Maksimum İvme		Maks.(A^+)		Maks.(A^-)
Ortalama İvme		Ort.(A^+)		Ort.(A^-)
Net Genlik, Süre Oranı	$\frac{(\sum(D^+)-\sum(D^-))\omega}{\eta(D)}$			

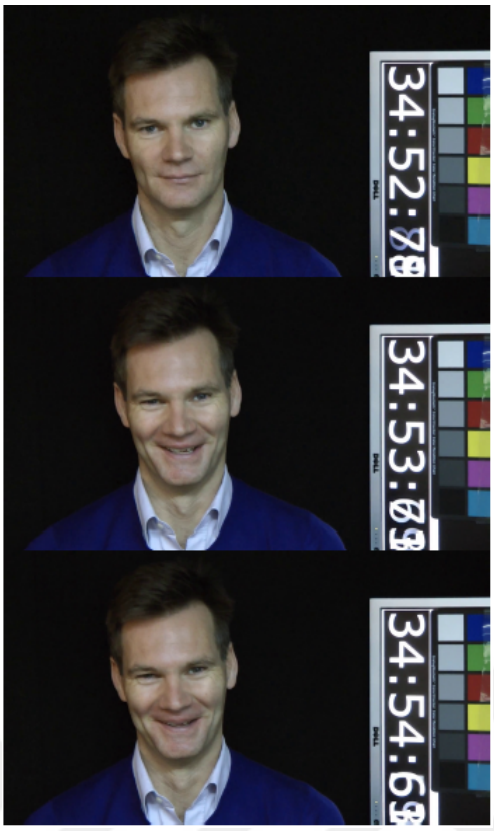
2.1.3.1 UvA-Nemo Gülümseme Veri Kümesi

UvA-NEMO veri kümesii [2, 62], 400 denekten toplanan 1240 adet yüksek çözünürlüklü (1920×1080) gülümseme videosundan oluşmaktadır. Deneklerin yaş aralığı 8 ila 76 arasında değişmekte olup, deneklerin 185'i kadın ve 215'i erkektir. Veri tabanında 597 adet spontan ve 643 adet kasıtlı gülümseme videosu bulunmaktadır. Deneklerden spontan gülümseme videoları elde edebilmek amacıyla, deneklere komik video klipler gösterilmiştir. Gülümseme videolarının ortalama süresi 3,9 saniyedir. UvA-NEMO veri tabanındaki bir gülümseme videosundan elde edilen örnek kareler Şekil 2.4'te verilmiştir.

2.1.3.2 Yüz Doğrulama Deneyleri

Dinamik özniteliklerin çıkartılması işlemi tamamlandıktan sonra, UvA-NEMO veri kümesinde bulunan 1240 adet gülümseme videosu için 648 boyutlu öznitelik vektörleri elde edilmiştir. Daha sonra elde edilen bu 1240×648 boyutlu öznitelik matrisinde her bir öznitelik sütunu için normalizasyon işlemi gerçekleştirilmiş ve her sütundaki değerler [0-1] değerleri arasında ölçeklendirilmiştir. Daha sonra elde edilen öznitelikler arasındaki Öklid uzaklığına bakılarak, elde edilen öznitelik vektörlerinin kişinin kimliği ile ilgili bilgi taşıyıp taşımadığı kontrol edilmiştir.

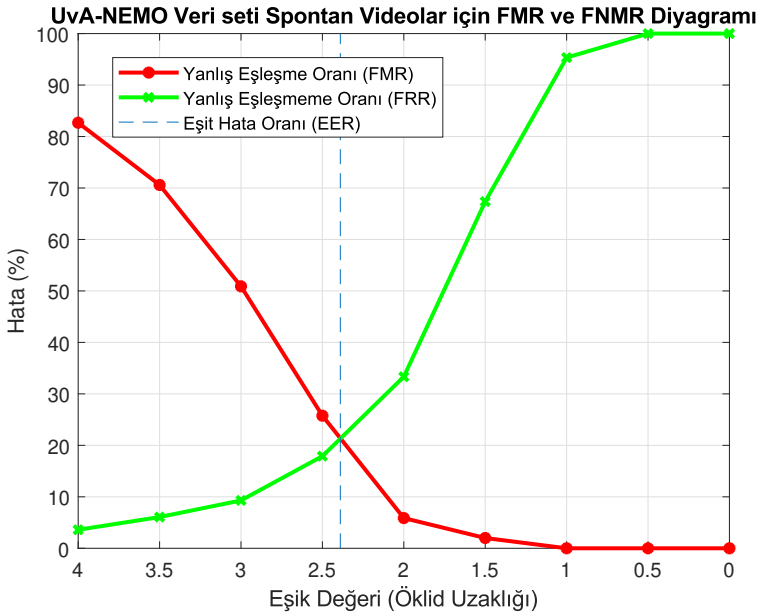
Bu amaçla, öncelikle UvA-Nemo gülümseme veri kümesinde bulunan videolar spontan (597 video) ve kasıtlı (643 video) videolar olarak ikiye ayrılmıştır. Daha sonra



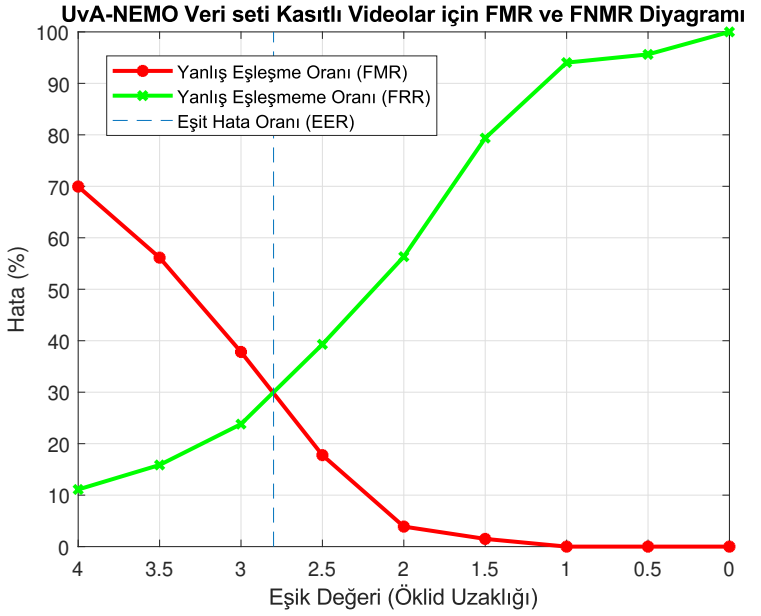
Şekil 2.4 UvA-NEMO veri kümesindeki bir gülümseme videosundan elde edilen örnek kareler.

videolar kişilerin kimliğine göre etiketlenmiş ve her bir videonun diğer videolar ile arasındaki Öklid uzaklığı değeri hesaplanmıştır. Daha sonra bu uzaklıklar kullanılarak, biyometrik sistemlerde kullanılan ölçütler olan Yanlış Eşleşme Oranı (False Match Rate (FMR)), Yanlış Eşleşmeme Oranı (False Non-Match Rate (FNMR) ve Eşit Hata Oranı (Equal Error Rate (EER)) değerleri hesaplanmıştır. Spontan videolar için hesaplanan farklı eşik değerlerine göre FAR ve FRR değerlerinin değişimi Şekil 2.5'te verilmiştir. Elde edilen FMR-FNMR grafiği incelendiği zaman, EER'nin elde edildiği Öklid uzaklığı değerinin 2.39 olduğu ve EER değerinin %21.25 olduğu gözlemlenmiştir. Kasıtlı gülmelerin bulunduğu videolar için elde edilen FAR, FRR ve EER sonuçları ise Şekil 2.6'da verilmiştir. Kasıtlı videolar için ERR'nin elde edildiği Öklid uzaklığı değerinin 2.8 olduğu ve EER değerinin %30 olduğu grafikte görülmektedir.

Deneysel çalışmaların ikinci aşamasında, videoların kasıtlı mı yoksa spontan mı olduğuna bakılmaksızın bütün videolar kullanılarak deneysel çalışmalar gerçekleştirilmiş ve FMR, FNMR ve EER değerleri elde edilmiştir. UvA-Nemo gülümseme veri tabanındaki bütün videoların kullanılması ile elde edilen FMR-FNMR diyagramı Şekil 2.7'te verilmiştir. Şekil 2.7'deki grafik incelendiği zaman, EER değerinin elde edildiği Öklid uzaklığının 2.8 olduğu ve EER değerinin %31.20 olduğu gözlemlenmiştir. Elde edilen üç adet FMR-FNMR diyagramları incelendiği zaman,

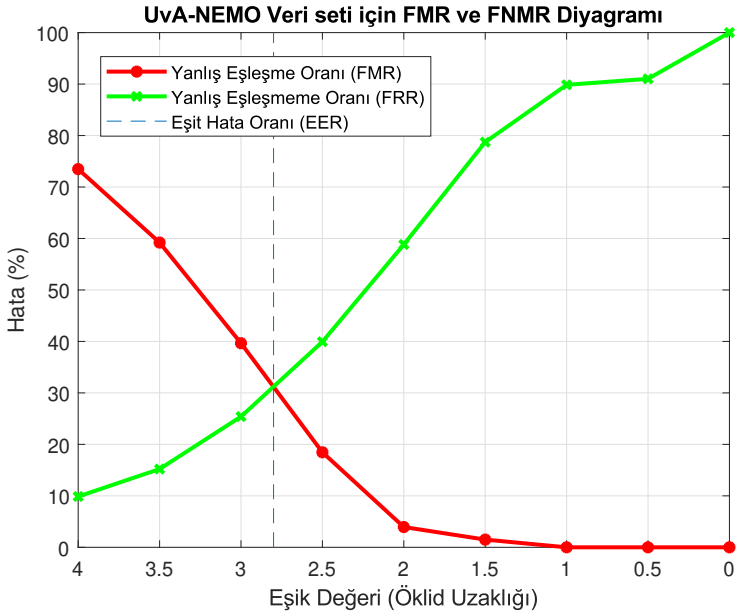


Şekil 2.5 Yalnızca spontan videolar için FMR ve FNMR diyagramı.



Şekil 2.6 Yalnızca kasıtlı videolar için FMR ve FNMR diyagramı.

en yüksek yüz doğrulama performansının, spontan şekilde gülme duygusunun kayıt edildiği videolardan elde edildiği gözlemlenmiştir. Bu da spontan şekilde alınan yüz videolarının kişinin kimliğiyle ilgili daha çok bilgi taşıdığı sonucuna varılmasını sağlamaktadır. Tüm videoların kullanıldığı deneysel çalışmada elde edilen yüz doğrulama performansının kalan iki deneme düzenlemesine oranla daha düşük olması, bu durumu destekler niteliktedir.



Şekil 2.7 Bütün (spontan ve kasıtlı) videolar için FMR ve FNMR diyagramı.

2.1.3.3 Öznitelik Seçim Algoritmaları ile Yüz Doğrulama Deneyleri

İlk iki deneysel çalışmalar sonucunda elde edilen FMR ve FNMR değerleri, yüze ait dinamiklerin kişilerin kimlik bilgilerini taşıdıklarını açık bir şekilde göstermiştir. Bununla birlikte, her bir videodan elde edilen 648 boyutlu öznitelik vektörünün hangi bölümlerinin kişiye ait bilgi taşıdığını anlamak ve öznitelik vektörünün boyutunu azaltmak amacı ile bu tez çalışması kapsamında bir öznitelik seçim algortiması kullanımaya karar verilmiştir. Bu tez çalışmasında, öznitelik seçim algortimasının kullanım amacı hem yüz tanıma başarımını arttırmak hem de daha az boyutlu bir öznitelik vektörü ile temsili sağlayarak önerilen methodun işlem yükünü azaltmaktır.

Öznitelik Önemlendirme (Aşırı Rastgele Ağaçlar Sınıflandırıcı) Algoritması ile Öznitelik Seçimi

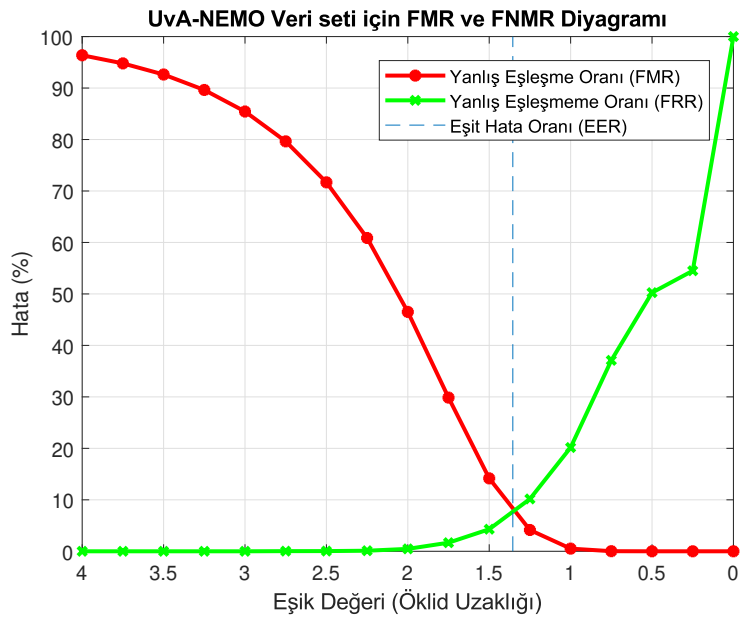
Bu tez çalışması kapsamında, 648 boyutlu istatistiksel yüz dinamikleri öznitelik vektöründen, öznitelik seçimi gerçekleştirmek amacı ile Aşırı Rastgele Ağaçlar Sınıflandırıcısı (ARAS) [257] kullanılmıştır. Aşırı Rastgele Ağaçlar Sınıflandırıcısı, bir sınıflandırma sonucu elde etmek için bir "ormanda" toplanan çok sayıda

bağıntılı karar ağacının sonuçlarını kullanan toplu bir öğrenme tekniğidir. Ekstra Ağaçlar Ormanındaki her Karar Ağacı, orijinal eğitim örneklerinin bir alt kümesini kullanmaktadır. Ağaçtaki her düğümde, her bir öznitelik kümesinden sınıflandırma için en iyi öznitelikleri seçmek amacıyla bir matematiksel kriter (Gini İndeksi gibi) kullanılmaktadır. Bu durum, birbirine bağlı birçok karar ağacının oluşturulmasına neden olmaktadır. Bu orman yapısını kullanarak öznitelik seçimi yapabilmek için, her bir öznitelik için matematiksel kriterin normalize edilmiş toplam indirgemesi hesaplanır. Matematiksel kriter olarak Gini İndeksi kullanılırsa bu hesaplama "Gini Önemi" olarak adlandırılır.

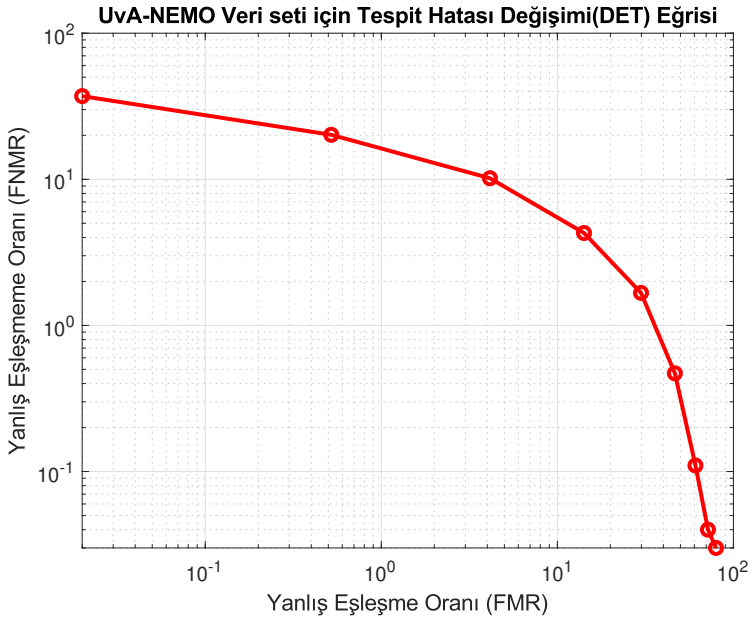
İndirgenmiş Öznitelik Vektörleri Kullanılarak Yüz Doğrulama Deneyleri

Bu çalışmada, öznitelik seçimi için "Gini Önemi" hesaplaması kullanılmıştır. Öznitelik seçimi sırasında, her bir öznitelik için hesaplanan Gini Önemi azalan düzende sıralanır ve ilk k adet öznitelik, öznitelik vektörünü elde etmek amacıyla seçilmektedir. Bu çalışmada da, yüz tanıma için her bir videodan elde edilmiş olan 648 boyutlu öznitelik vektörü içerisinden Gini Önemi en yüksek (yani 0,002'den büyük) olan 128 adet özniteliği seçilerek yeni bir tanımlayıcı öznitelik vektörü oluşturulmuştur. Yeni öznitelik vektörleri oluşturulduktan sonra, elde edilen bu 1240 × 128 boyutlu öznitelik matrisinde her bir öznitelik sütunu için normalizasyon işlemi gerçekleştirilmiş ve her sütundaki değerler [0-1] değerleri arasında ölçeklendirilmiştir. Daha sonra elde edilen öznitelikler arasındaki Öklid uzaklığına bakılarak, elde edilen öznitelik vektörlerinin kişinin kimliği ile ilgili bilgi taşıyıp taşımadığı kontrol edilmiştir. Bu amaç doğrultusunda, videolar kişilerin kimliğine göre etiketlenmiş ve her bir videonun diğer videolar ile arasındaki Öklid uzaklığı değeri hesaplanmıştır. Daha sonra bu uzaklıklar kullanılarak, biyometrik sistemlerde kullanılan ölçütler olan Yanlış Eşleşme Oranı (FMR), Yanlış Eşleşmeme Oranı (FNMR) ve Eşit Hata Oranı (EER) değerleri hesaplanmıştır. Elde edilen FMR-FNMR diyagramı ve Tespit Hatası Değişimi (DET) eğrisi sırasıyla Şekil 2.8 ve Şekil 2.9'te verilmiştir.

Şekil 2.8'te verilen FMR-FNMR diyagramında, Eşit hata oranı (EER)'nın, FMR'nin FNMR'ye eşit olduğu noktada kesikli bir çizgiyle işaretlendiği görülmektedir. Öznitelik seçimi algoritmsı kullanılmadan önce %31.20 olarak elde edilen EER değerinin, öznitelik seçimi algoritması kullanıdıktan sonra %7,42'ye düşdüğü gözlemlenmiştir. Bu sonuç, kullanılan öznitelik seçimi algoritması ile, gülümseme aktivitesi sırasında elde edilen yüz dinamiği bilgilerinin kullanarak yüz doğrulama performansında önemli bir artış sağladığını göstermiştir. Gerçekleştirilen deneysel çalışmalar sonucunda, gülümseme aktivitesi sırasında elde edilen yüze ait dinamik özniteliklerin, kişinin kimliğine ait bilgiler taşıdığı, özellikle öznitelik seçim algoritması kullanıldıktan sonra, net bir şekilde gösterilmiştir.



Şekil 2.8 UvA-NEMO gülümseme veri kümesinde, 128 boyutlu dinamik öznitelik vektörleri kullanılarak gerçekleştirilen yüz doğrulama deneyleri için FMR-FNMR diyagramı.



Şekil 2.9 UvA-NEMO gülümseme veri kümesinde, 128 boyutlu dinamik öznitelik vektörleri kullanılarak gerçekleştirilen yüz doğrulama deneyleri için Tespit Hatası Değişimi (DET) eğrisi.

3
YÜZ TANIMA İÇİN HİBRİT BİR YÖNTEM

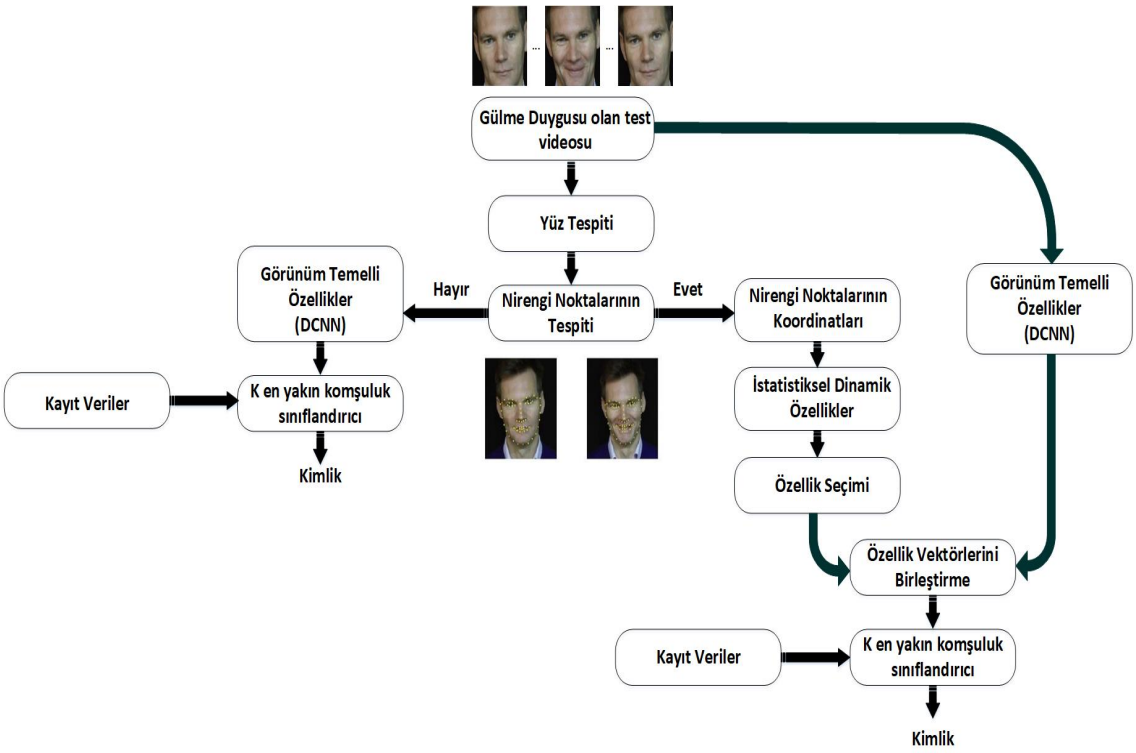
Bu bölümde, DESA'dan elde edilen görünüme dayalı ana biyometrik veriler ile Bölüm 2'de açıklanan dinamik özniteliklerden elde edilen yumuşak biyometrik verileri birlikte kullanan hibrit bir yüz tanıma sistemi (HYT) önerilmektedir. Bu bölümün alt başlıklarında, önerilen hibrit yüz tanıma sisteminin farklı zorlu koşullar altındaki yüz tanıma başarımını incelemek amacıyla, iki farklı veri kümesi kullanılarak gerçekleştirilen deneysel çalışmalar ve bu çalışmaların sonuçlarından bahsedilecektir.

3.1 Hibrit Yüz Tanıma Sistemi

Önerilen hibrit yüz tanıma sisteminin (HYT) blok diyagramı Şekil 3.1'de verilmiştir. İlk olarak, gülümseme ifadesi içeren belirli bir videodan elde edilen her bir video karesinde yüzün konumu tespit edilecektir. Ardından, yüze ait dinamikleri elde etmek amacıyla gözler, kaşlar, burun, dudak ve çene çevresinde yüz nirengi noktaları tespit edilecektir. Yüzdeki nirengi noktaları başarılı bir şekilde algılanabiliyorsa, nirengi noktalarından çıkarılan dinamik tabanlı öznitelikler ve önceden eğitilmiş bir derin evrişimli sinir ağından (DESA) elde edilen görünüm tabanlı öznitelikler, olumsuz koşullar altında performansı artırmak için yüz tanıma işlemi için birlikte kullanılacaktır. Yüzdeki nirengi noktaları başarılı bir şekilde algılanamadığı koşullarda, yalnızca görünüm tabanlı öznitelik vektörleri kullanılarak yüz tanıma gerçekleştirilecektir.

3.1.1 Yüz Tespiti

Önerilen hibrit yüz tanıma sisteminin ilk adımı, gülümseme ifadesi içeren videonun tüm karelerindeki yüz konumlarını tespit etmektir. Literatürde yüz tanıma işlemini gerçekleştirmek için çok sayıda yöntem önerilmiştir. Literatürde yüz tespiti için yaygın kullanılan yöntemlerden biri, Haar kaskadlarını kullanan Viola-Jones (VJ) yüz algılayıcıdır [64]. VJ yüz algılayıcı ön yüzleri algılamada başarılı olmasına rağmen, farklı kafa duruşları için aynı performansı elde edememektedir. 2005 yılında Dalal ve



Şekil 3.1 Önerilen hibrit yüz tanıma sisteminin blok diyagramı.

Triggs, VJ yüz dedektörüne kıyasla daha başarılı sonuçlar elde eden, Yönlendirilmiş Gradyanların Histogramı (HOG) ve Destek Vektör Makinesi (SVM) sınıflandırıcısına [258] dayandığı başka bir yüz tespit yöntemi önermişlerdir. Son yıllarda, derin öğrenme ağlarının kullanımı ile birlikte, farklı koşul ve durumlarda yüz tespiti gerçekleştirebilen yöntemlerin ortaya çıkmıştır. Tek Vuruş Çoklu-Kutu Dedektörü (SSD) yöntemi [75], derin öğrenme ağı mimarisi olarak ResNet-10 mimarisini kullanmaktadır. Eğitim verisi olarak, web'den toplanan ve farklı açı ve mesafelerde yüz verileri kullanılmıştır. Yüz tespiti için kullanılan bir başka DESA tabanlı yöntem ise, Maksimum Marjin Nesne Tespiti'dir(MMOD) [259]. Geleneksel DESA mimarilerinin eğitimi, büyük ölçekli veri tabanları gerektirirken, MMOD ağının eğitimi, farklı yüz veritabanlarından alınan yaklaşık 7000 yüz görüntüsü kullanılarak gerçekleştirilmiş ve bu veri miktarı ağın eğitimi için yeterli olmuştur.

Bu tez çalışmasında, çeşitli kafa duruşlarında daha doğru sonuçlar verdiği ve test edilen diğer yüz tanıma yöntemlerine kıyasla çok daha hızlı çalıştığı için yüz tespiti işlemi için MMOD ağı tercih edilmiştir.

3.1.2 Yüz Nirengi Noktalarının Tespit Edilmesi

Bu tez çalışmasında, Bölüm 2.1.1 "Yüz Nirengi Noktalarının Tespit Edilmesi" alt başlığı altında belirtildiği üzere, yüz nirengi noktalarını tespit etmek amacıyla Kazemi ve

arkadaşları tarafından önerilen yöntem kullanılmaktadır [79]. Kazemi ve arkadaşları tarafından önerilen yöntem, kare hata kaybının toplamını optimize eden ve doğal olarak eksik verileri yöneten bir regresyon ağaçları topluluğunu öğrenmek için gradyan artırmaya dayalı bir yöntemdir.

3.1.3 Görünüm-Temelli Özniteliklerin Elde Edilmesi

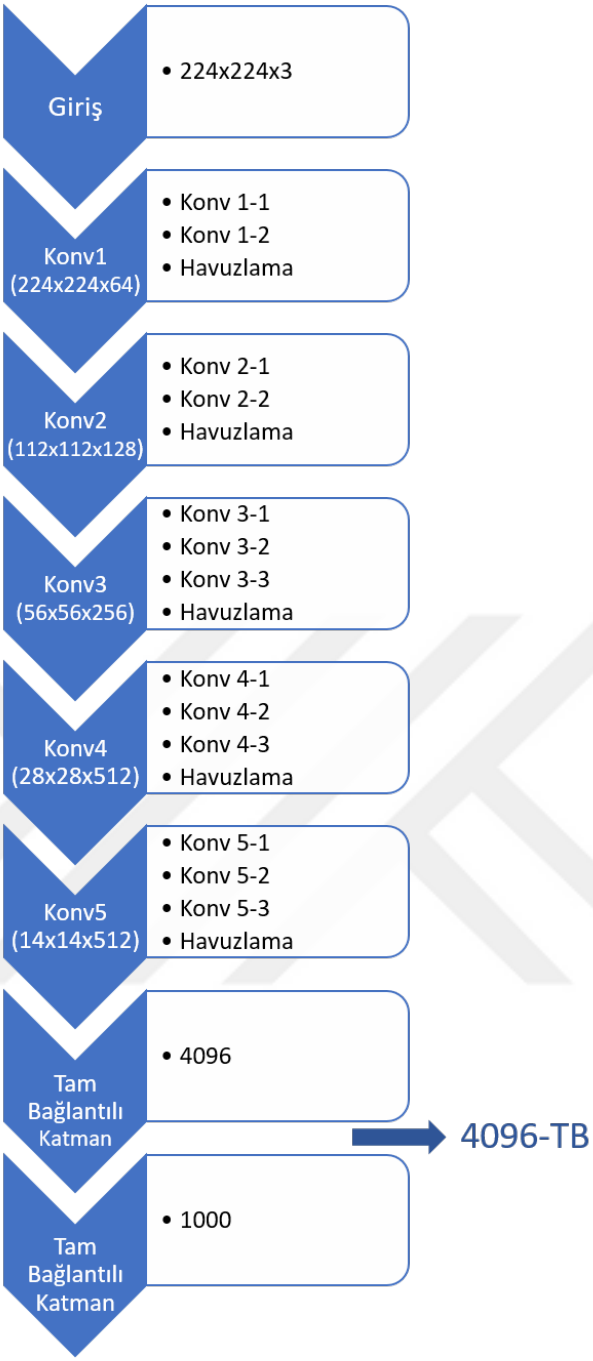
Yüz tanıma sistemlerinin performansı, derin sinir ağlarının kullanılmasıyla birlikte hızla artmış ve büyük ölçekli veri tabanlarında %99'a varan yüz tanıma doğruluğu sağlamıştır [260–262]. Bu tez çalışmasında, olumsuz koşullar altında görünüm tabanlı öznitelik vektörlerini çıkarmak için önceden eğitilmiş üç farklı derin sinir ağı mimarisini (VGG-Face [37], VGG-Face2 [139] ve ArcFace [263]) kullanılacaktır. Bu üç farklı derin sinir ağı mimarisi hakkında detaylı bilgiler aşağıda verilmiştir.

3.1.3.1 VGG-Face

VGG-Face [37], yüz tanıma gerçekleştirmek amacıyla 2015 yılında Parkhi ve arkadaşları tarafından VGG-Çok Derin-16 derin öğrenme ağı mimarisine dayanan ve yaklaşık 2,600 denekten elde edilen yaklaşık 2,6 milyon yüz görüntüsü ile eğitilerek oluşturulmuştur. VGG-Face'in oluşturulmasında kullanılan VGG-Çok Derin-16 derin öğrenme ağı mimarisi Şekil 3.2'de verilmiştir. Labeled Faces in the Wild [120] ve Youtube Faces Databases [152] üzerinde gerçekleştirilen deneysel çalışmalar sonucunda, sırasıyla %97,27 ve %92,8 yüz tanıma doğruluğu elde edilmiştir. Bu ağa verilecek olan giriş görüntüsü 224×224×3 boyutundadır ve ağın çıkışında elde edilen öznitelik vektörünün boyutu, tam bağlı katmanlardan hemen önce fc7 katmanından elde edilen 4096×1 boyutlu vektördür. Bu tez çalışmasında, görünüm temelli öznitelik vektörü olarak bu 4096×1 boyutlu vektör kullanılmıştır.

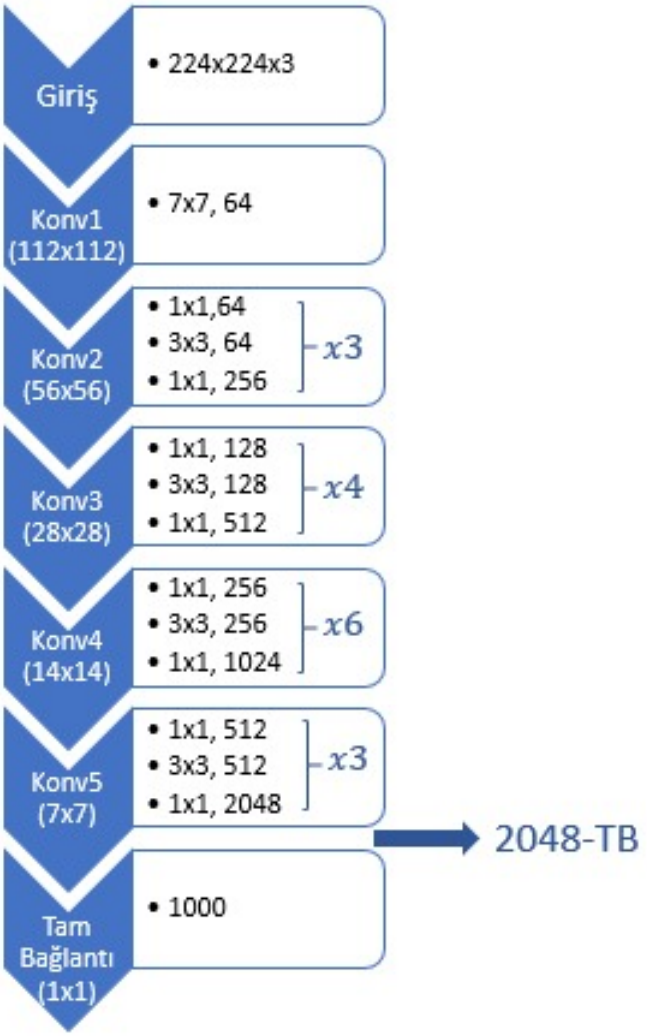
3.1.3.2 VGG-Face2

2018 yılında, Cao ve arkadaşları tarafından yüz tanıma için, 9131 denekten farklı ışıklandırma ve poz değişimleri altında elde edilen yaklaşık 3.31 milyon yüz görüntüsünü içeren VGG-Face2 [139] adlı büyük ölçekli bir veri tabanı toplanmıştır. ResNet-50 ve Squeeze-and-Excitation bloklarına sahip Resnet-50 mimarileri [217, 218], VGG-Face2 veri kümesi kullanılarak eğitilmiştir ve elde edilen modelin yüz tanıma performansının VGG-Face'den daha iyi olduğu gözlemlenmiştir. Bu ağa verilecek olan giriş görüntüsü 224×224×3 boyutundadır ve ağın çıkışında elde edilen öznitelik vektörünün boyutu, tam bağlı katmanlardan hemen önce fc7 katmanından elde edilen 2048×1 boyutlu vektördür. Bu tez çalışmasında, ağ mimarisi olarak ResNet-50 mimarisi tercih edilmiştir ve görünüm temelli öznitelik vektörü olarak



Şekil 3.2 VGG-16 derin sinir ağı mimarisi. [264]

2048×1 boyutlu vektör kullanılmıştır. ResNet-50 mimarisinin detaylı gösterimi Şekil 3.3'de verilmiştir.



Şekil 3.3 ResNet-50 derin sinir ağı mimarisi. [217]

3.1.3.3 ArcFace

2019'un başlarında Deng ve arkadaşları tarafından, yüz tanıma için üst düzey ayırt edici öznitelikler elde etmek amacıyla toplamalı açısal marj kaybına (ArcFace) [263] dayalı bir yöntem önerilmiştir. 10 farklı geniş kapsamlı yüz veri tabanını kullanılarak gerçekleştirilen deneysel sonuçlar, ArcFace'in yüz tanıma için çok yüksek test performansı sağladığını göstermiştir. Bu tez çalışmasında, ArcFace için seçilen derin öğrenme mimarisi Inception-ResNet50 derin öğrenme mimarisidir. ArcFace ağlarının eğitimi MS-Celeb-1M[136] ve VGG-Face2 veri tabanları kullanılarak gerçekleştirilmektedir. Ağın eğitimi tamamlandıktan sonra, fc1 katmanından elde

edilen 512 boyutlu öznitelik vektörü bu çalışmada, görünüm-temelli öznitelik vektörü olarak kullanılmıştır.

3.1.4 İstatistiksel Dinamik Yüz Özniteliklerinin Elde Edilmesi

Yüze ait dinamiklerin kullanılarak yüz tanıma gerçekleştirmek için önerilen yöntem, Şekil 2.1'de gösterilmiştir. Bu amaç doğrultusunda gerçekleştirilecek işlemler, detaylı bir şekilde Bölüm 2.1.2 "Yüze Ait Dinamik Özniteliklerin Çıkarılması" kısmında anlatılmıştır. Bu bölümde bahsedilen işlemlerin gerçekleştirilmesi sonucunda, veri kümesinde bulunan her bir videodan 648 boyutlu öznitelik vektörleri elde edilmiştir. Daha sonraki bölümde, Bölüm 2.1.3.3 "Öznitelik Seçim Algoritmaları ile Yüz Doğrulama Deneyleri" bahsedildiği üzere, her bir videodan elde edilen 648 boyutlu öznitelik vektörü öznitelik seçim algoritması (Aşırı Rastgele Ağaçlar Sınıflandırıcı) yardımıyla 128 boyutlu öznitelik vektörüne dönüştürülmüştür. Yine aynı bölümde bahsedilen deneysel çalışmaların sonuçları, elde edilen öznitelik vektörünün kişilerin kimliğine ait bilgi taşıdığını göstermiştir.

3.1.5 Sınıflandırma

DESA'dan görünüm-temelli öznitelik vektörünün ve gülümseme aktivitesi sırasında oluşan yüze ait istatistiksel dinamik öznitelik vektörünün elde edilmesinden sonra, bu öznitelik vektörlerinin birleştirilmesi işlemi gerçekleştirilir. Daha sonra ise elde edilen bu öznitelik vektörü, kayıtlı veriler arasında bulunan öznitelik vektörleri ile benzelikleri incelenecek ve kişi tanıma işlemi için sınıflandırma işlemi gerçekleştirilecektir. Bu tez çalışmasında, sınıflandırıcı olarak k-en yakın komşuluk algotirmasının kullanılmasına karar verilmiştir. K-en yakın komşuluk (k-NN), her giriş verisinin uzaklık benzerliklerine göre rastgele seçilen komşuluk (k) sınıfına atandığı bir makine öğrenmesi yöntemidir [265]. Algoritma ilk olarak 1952 yılında Evelyn Fix ve Joseph Hodges tarafından önerilmiştir [266]. Daha sonra önerilen yöntem farklı araştırmacılar tarafından geliştirilmiştir. K-en yakın komşuluk algoritmasının hesaplama kolaylığı, kullanılması gereken parametrelerin azlığı ve gürültüye karşı dayanıklılık gibi avantajları bulunmaktayken, aynı zamanda bazı dezavantajları bulunmaktadır. Bu dezavantajlardan ilki "k" değerinin belirlenmesinin önemli olmasıdır. K değerinin çok küçük ya da büyük seçilmesi sistemin performansını olumsuz yönde etkileyebilmektedir. Bunun yanı sıra, Öklid uzaklığının karmaşık verileri ayırmakta zorluk çekmesi ve yüksek boyutlu verilerin sınıflandırılmasında yavaşlama gibi durumlar, algoritmanın diğer olumsuz yönleri olarak sayılabilmektedir. Bu tez çalışmasında, yapılan ilk denemelerde k-en yakın komşuluk algoritmasının başarılı sonuçlar elde edilmesi nedeni ile sınıflandırıcı olarak bu algoritma seçilmiştir.

"k" sayısına karar vermek amacıyla, gerçekleştirilen deneysel çalışmalar sırasında k değeri 1 ila 8 arasında değişen sayılar seçilmiştir ve elde edilen sonuçlar doğrultusunda hibrit yüz tanıma sistemi için k değeri 7 olarak belirlenmiştir.

3.2 Hibrit Yüz Tanıma Sistemini Değerlendirmek için Gerçekleştirilen Deneysel Çalışmalar

Bu bölümde ilk olarak, deneysel çalışmalarda kullanılan video veri kümeleri hakkında detaylı bilgiler verilmiştir. Ardından bu çalışmalar sırasında kullanılan değerlendirme yöntemi ve test edilen bozulmaların (olumsuz koşullar) neler olduğu açıklanmıştır. Daha sonraki bölümde, farklı koşullar altında, sadece gülümseme aktivitesi sırasında elde edilen istatistiksel dinamik öznitelikler kullanılarak kişi tanımadan elde edilen sonuçlar, sadece DESA'lardan elde edilen görünüm temelli öznitelik vektörleri kullanılarak kişi tanıma gerçekleştirilen çalışmalardan elde edilen sonuçlar ve iki öznitelik vektörünün birleşmesi ile önerilen hibrit yüz tanıma sisteminden elde edilen sonuçlar ve bu sonuçların yorumlanması sırası ile verilmiştir.

3.2.1 Video Veri Kümeleri

Literatürde, aynı denekten birden fazla gülümseme videosunu içeren yalnızca birkaç adet kamuya açık veri kümesi bulunmaktadır. Farklı yüz ifadelerini ve duygularını içeren video veri kümeleri genel olarak duygu tanıma çalışmaları için toplanmaktadır ve bu nedenle video veri kümelerinde, aynı deneğin aynı duygusuna ait az sayıda video bulunmaktadır. Bu tez çalışmasında, önerilen HYT sistemini test etmek amacıyla, her denek için en az iki gülümseme videosu içeren iki farklı video veri kümesi kullanılmıştır.

UvA-NEMO veri kümesi [2, 62], 400 denekten toplanan 1240 adet yüksek çözünürlüklü (1920x1080) gülümseme videosundan oluşmaktadır. Deneklerin yaş aralığı 8 ila 76 arasında değişmekte olup, deneklerin 185'i kadın ve 215'i erkektir. Veri kümesinde 597 adet spontan ve 643 adet kasıtlı gülümseme videosu bulunmaktadır. Deneklerden spontan gülümseme videoları elde edebilmek amacıyla, deneklere komik video klipler gösterilmiştir. Gülümseme videolarının ortalama süresi 3,9 saniyedir. Bu çalışma kapsamında, gülümseme türünden bağımsız olarak bir deneğe ait bütün gülümseme videolarını kullanılmıştır.

Deneysel çalışmalarda kullanılan ikinci veri kümesi, yaşları 23-38 arasında değişen 18 denekten 378 video içeren FEEDTUM yüz ifadesi veri kümesidir [267]. Her denek, altı farklı yüz ifadesi ve nötr durumu içeren 21 videoya sahiptir. Önerilen yöntemi test

etmek için sadece gülümseme ifadesine ihtiyacımız olduğundan, 18 denekten elde edilen 54 gülümseme videosu kullanılmıştır.

Deneysel çalışmalarda kullanılan video veri kümeleri hakkında detaylı bilgiler, Tablo 3.1'de verilmiştir.

Tablo 3.1 Kullanılan video veri kümelerinin açıklaması.

	Tanımlamalar						
Video Veri Kümeleri	**Video Sayısı**	**Denek Sayısı**	**Cinsiyet**	**Yaş Aralığı**	**Çözünürlük**	**Video Süresi (s)**	**Sks**
UvA-NEMO Veri kümesi	1240	400	185 kadın, 215 erkek	8 ila 76 arası	1920 x 1080	ort. 3.9	50
FEEDTUM Veri kümesi	378	18	9 kadın, 9 erkek	23 ila 38 arası	640 x 480	ort. 3	25

3.2.2 Değerlendirme Metotları ve Test Edilen Bozulmalar

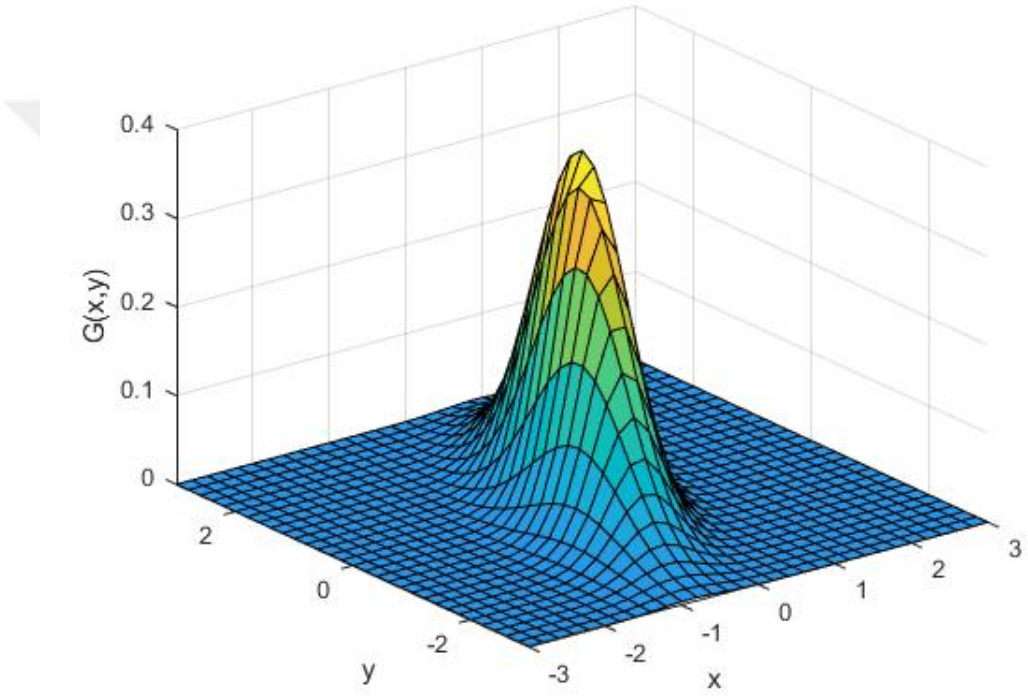
Bu tez çalışmanın temel amacı, gülümseme ifadesinden elde edilen istatistiksel dinamik özniteliklerin, olumsuz koşullar altında yüz tanıma performansı düşen derin öğrenme tabanlı yüz tanıma sistemlerinin performansını iyileştirme potansiyeline sahip olduğunu göstermektir.Bu nedenle, öncelikle daha önceden eğitilmiş olan derin sinir ağlarının yüz tanıma performansını olumsuz koşullar altında değerlendirilmesi gerçekleştirilmiştir. Ardından, önceden eğitilmiş DESA'lardan elde edilen görünüm tabanlı öznitelikler ile istatistiksel dinamik özniteliklerin birleştirilmesinin olumsuz koşullar altında yüz tanıma performansı üzerindeki etkisi araştırılmıştır.

3.2.2.1 Gauss Bulanıklığı

Gauss filtreleme, görüntüleri bulanıklaştırmak ve gürültüyü, ayrıntıları kaldırmak amacıyla kullanılmaktadır [268]. Bu anlamda, ortalama filtreye benzemekle beraber, Gauss (çan şeklinde) kamburunun şeklini temsil eden farklı bir çekirdek kullanmaktadır. Bu çekirdeğin bazı özel özellikleri vardır. Gauss fonksiyonu imge bulanıklaştırma için iki boyutlu olarak kullanılmaktadır. Gauss filtreleme için kullanılan denklem aşağıda verilmektedir;

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2.y^2}{2\sigma^2}} \tag{3.1}$$

Yukarıdaki denklemde σ dağılımın standart sapmasıdır. Ayrıca, dağılımın ortalama değerinin (0,0) olduğunu varsayılmıştır. σ değerinin 1, dağılımın ortalama değerinin (0,0) olduğu Gauss dağılımı Şekil 3.4'de gösterilmiştir. Gözetleme kameraları, odak dışında kalma veya hareketden kaynaklı bulanıklık nedeniyle bulanık videolar çekebilmektedir. Odak dışı bulanıklığı simüle etmek ve bulanıklık bozulmalarının derin öğrenme tabanlı yüz tanıma yöntemleri üzerindeki etkisini araştırmak amacıyla test görüntülerine farklı standart sapmalara sahip Gauss düşük geçiş filtreleri uygulandı. Bu tez çalışmasında, Gauss filtresi için kullanılan standart sapma (σ) değeri 2,5 ile 15 arasında değişmektedir.



Şekil 3.4 Ortalama (0,0) ve σ=1 için Gauss dağılımı.

3.2.2.2 Aydınlatma Değişimleri

Aydınlatma seviyelerindeki değişimlerin derin öğrenme ağındaki etkisini incelemek amacıyla, veri kümesindeki resimlerin aydınlatma seviyeleri değiştirilmiştir. Aydınlatma seviyesindeki değişikliklerin yüz tanıma performansı üzerindeki etkilerini araştırmak için iki farklı yöntem uygulanmıştır. İlk yöntemde, görüntülerdeki aydınlatma seviyeleri, piksel değerlerine 5 ile 200 arasında değişen sabit bir değer eklenerek değiştirilmiştir. İkinci yöntemde piksel yoğunlukları 1 ile 5 arasında değişen sabit değerlerle çarpılmıştır. Görüntülerin piksellerin değerleri, her işlemden sonra [0-255] aralığına ölçeklenmiştir.

3.2.2.3 Gauss Gürültüsü Ekleme

Resimlere Gauss gürültüsü ekleme işlemi, belirlenen bir sigma değerine göre Gauss dağılımı oluşturup, bu dağılıma göre piksel değerlerinde yapılan değişimlerle sağlanmaktadır [269]. Gerçekleştirilen deneysel çalışmalarda, test görüntülerine ek olarak Gauss gürültüsü ilave edilmiştir. Eklenen Gauss gürültüsünde ortalama değer 0 olarak alınmış ve standart sapma değerleri 0.1 ila 0.5 arasında değişmiştir.

3.2.2.4 Tuz ve Biber Gürültüsü Ekleme

Tuz ve biber gürültüsü, uygulandığı resimlere nokta nokta olarak görülecek şekilde piksel değerlerinin değiştirilmesine neden olmaktadır [270]. Uygulanan işleme tuz ve biber gürültüsü adının verilmesinin sebebi, resimlere eklenen siyah ve beyaz noktalardan almaktadır. Tuz ve biber gürültüsü eklenmesi işlemi, önceden belirlenen orana göre resimdeki rastgele seçilmiş piksellerin değerlerinin 0 ve 255 ile değiştirilmesi sayesinde gerçekleştirilir. Bu tez çalışmasında, tuz ve biber gürültüsü için belirlenen oran değerleri 0.1 ila 0.5 arasında değişmektedir.

3.2.2.5 Kontrast Değişimleri

Kontrast değişimleri, derin öğrenme ağındaki etkisini incelemek amacıyla veri kümesindeki resimlerin kontrastı değiştirilmiştir. Test görüntülerinin kontrastını değiştirmek için, belirlenen giriş alt ve üst sınır değerlerinin altında ve üstünde kalan piksel değerlerini yine belirlenen çıkış sınır değerlerine eşitlenir. Aradaki değerler de eğer gama değeri 1 seçildiyse lineer olarak hesaplanır. Bu çalışmada gama değeri 1 olarak seçilmiştir. Gerçekleştirilen deneysel çalışmalarda, en geniş giriş aralığı [0.1–0.9] ve en dar aralık [0.49–0.51] olarak belirlenmiştir.

3.2.3 Video Veri Kümeleri Kullanılarak Gerçekleştirilen Deneysel Çalışmalar

Bu tez çalışması kapsamında, kapalı küme yüz tanıma için önerilen HYT sisteminin performansını incelemek amacıyla, UvA-NEMO gülümseme veri kümesi ve FEEDTUM veri kümesi kullanılarak deneysel çalışmalar gerçekleştirilmiştir. Bu çalışmalar kapsamında, farklı koşullar altında (farklı ışıklandırma, bulanıklaştırma, gürültü vs.) yüze ait dinamikler kullanılarak elde edilen yüz tanıma başarımları, DESA temelli görünüm öznitelikleri kullanılarak elde edilen yüz tanıma başarımları ve yüze ait dinamiklerin ve görünüm temelli özniteliklerin birleştirilmesi ile elde edilen hibrit öznitelik vektörlerinin kullanılması ile elde edilen yüz tanıma başarımları sırasıyla incelenmiş ve alt başlıklarda detaylı bir şekilde anlatılmıştır.

3.2.3.1 İstatistiksel Dinamik Öznitelikler Kullanılarak Gerçekleştirilen Yüz Tanıma Deneyleri

Bu bölümde, istatistiksel dinamik öznitelikler kullanılarak yüz tanıma gerçekleştirmek amacıyla veri kümesinde bulunan videolar eğitim(kayıt verisi) ve test olarak ikiye ayrılmıştır. UvA-NEMO veri kümesinde toplam 1240 adet video bulunmaktadır. 30 deneğe ait tek bir video bulunması ve deneysel çalışmalar için biri eğitimde(kayıtta) diğeri testte olacak şekilde en az iki videoya ihtiyaç duyulması nedeni ile bu 30 kişi deneysel çalışmalara katılmamıştır. Geriye kalan 1210 adet video ise, veri kümesinde bulunan her bir deneğin video sayısı göz önünde bulundurularak, çoğunluk kayıt verisi olarak ayrılacak şekilde eğitim ve test verisi olarak ayrılmıştır. Bu işlem sonucunda 817 video kayıt verisi olarak kullanılırken, 393 video test verisi olarak ayrılmıştır. Videoların ayrılması işlemi gerçekleştirildikten sonra, test verilerine farklı koşulları simüle etmek amacıyla çeşitli işlemler uygulanmış ve k-en yakın komşuluk sınıflandırıcı kullanılarak deneysel çalışmalar gerçekleştirilmiştir. K değeri, bu çalışma için "1" olarak seçilmiştir. Bir kişiye ait en az iki adet video bulunması nedeniyle çapraz doğrulama sayısı 2 olarak belirlenmiştir. Gerçekleştirilen deneysel çalışmalar sonucunda elde edilen yüz tanıma performansları Tablo 3.2'te verilmiştir. Tablodaki sonuçlar incelendiği zaman, aydınlatma değişimlerinde ve kontrast değişimlerinin bulunduğu durumların, gürültü eklenen durumlara oranla daha stabil olduğu gözlemlenmiştir.

3.2.3.2 Görünüm Temelli ve Hibrit Öznitelikler Kullanılarak Gerçekleştirilen Yüz Tanıma Deneyleri

UvA-NEMO veri tabanında yüz tanıma performansını araştırmak için videonun uzunluğundan bağımsız olarak düzenli aralıklarla her videodan 12 kare çıkarılarak bir görüntü veri kümesi oluşturulmuştur. Bu görüntü veri kümesi oluşturulurken yalnızca bir videosu olan deneklerin videoları kullanılmamış olup, sonuç olarak 370 deneğe ait 1210 video kullanılarak veri kümesi oluşturulmuştur. Böylece 1210 videodan, 14.520 yüz görüntüsü içeren bir veri tabanı elde edilmiştir. Eğitim kümesi bu görüntülerin 9804'ünden oluşmakta olup, test için 4716 adet görüntü kullanılmıştır. Verilerin ayrılması işlemi gerçekleştirilirken, her deneğe ait toplam video sayısı dikkate alınmış ve her denek için eğitim ve test videosu olarak ayrılan videoların farklı olduğundan emin olunmuştur.

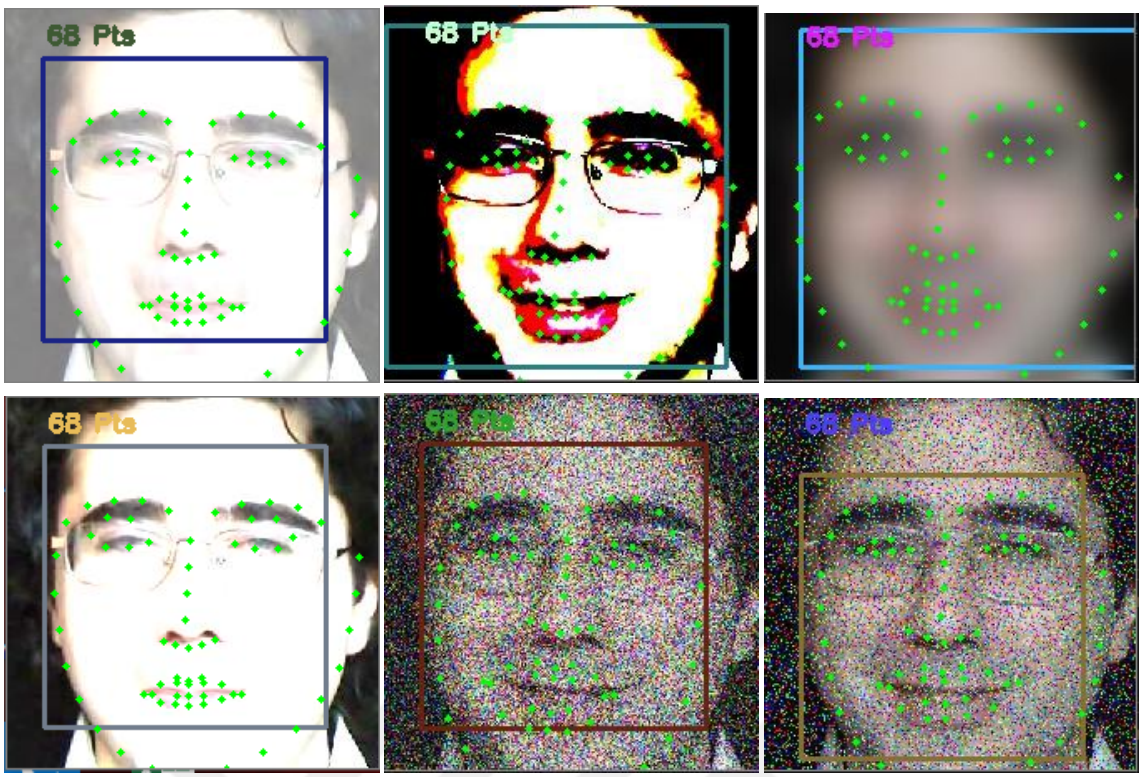
Deneylerin ilk bölümünde UvA-NEMO gülümseme video veri tabanından elde edilen test görüntülerine çeşitli distorsiyonlar uygulanmış ve bu görüntüler kullanılarak önceden eğitilmiş derin sinir ağlarının olumsuz koşullar altındaki yüz tanıma performansı araştırılmıştır.

Tablo 3.2 Çeşitli bozulmalar için yalnızca istatistiksel yüz dinamiklerini kullanılarak UvA-NEMO gülümseme veri kümesinden elde edilen yüz tanıma doğrulukları (%).

Görüntü Bozulmaları	**Parametreler / Başarım (%)**				
Bozulma Yok	81.02 ±0.16				
Gauss Bulanıklığı	**$\sigma = 2.5$**	**$\sigma = 5$**	**$\sigma = 7.5$**	**$\sigma = 10$**	
	76.20 ±0.57	52.14 ±0.36	24.33 ±0.46	17.65 ±0.74	
Toplamsal Aydınlatma	**5**	**10**	**15**	**20**	**25**
	81.02 ±0.16	81.02 ±0.16	81.02 ±0.16	81.02 ±0.16	81.02 ±0.16
Toplamsal Aydınlatma	**50**	**70**	**100**	**125**	
	81.02 ±0.16	81.02 ±0.16	78.07 ±0.80	62.84 ±0.28	
Çarpımsal Aydınlatma	**1**	**1.5**	**2**	**2.5**	
	81.02 ±0.16	76.20 ±0.20	69.79 ±0.29	42.25 ±0.25	
Gauss Gürültüsü	**$\sigma = 0.01$**	**$\sigma = 0.05$**	**$\sigma = 0.1$**		
	51.24 ±0.34	4.54 ±0.98	2.14 ±0.14		
Kontrast Değişimleri	**[0.1-0.9]**	**[0.2-0.8]**	**[0.3-0.7]**	**[0.4-0.6]**	**[0.49-0.51]**
	81.02 ±0.16	81.02 ±0.16	81.02 ±0.16	81.02 ±0.16	78.07 ±0.27
Tuz ve Biber Gürültüsü	**0.1**	**0.2**	**0.3**		
	36.90 ±0.40	7.22 ±0.32	2.14 ±0.04		

Deneylerin ikinci bölümünde, önceden eğitilmiş DESA'lardan elde edilen öznitelik vektörleri ile gülümseme videolarından elde edilen istatistiksel dinamik özniteliklerin birleştirilmesinin etkisi araştırılmıştır. Test edilen bozulmaların çoğu için yüz işaretleri tespit edilebilmektedir. Şekil 3.5'de, altı tür bozulma kategorisinin her biri için yüz nirengi noktalarının tespit edilebilecek olduğu en yüksek bozulma değerlerine sahip görüntüler gösterilmektedir. Şekil 3.5'den, görüntünün yapısı bozulmuş olsa da, nirengi noktalarının tespitinin, ciddi bozulmalar altında dinamik özniteliklerden yararlanmamızı sağladığı görülebilmektedir.

UvA-NEMO gülümseme veri tabanı için hem DESA'lardan elde edilen görünüme dayalı özniteliklerin hem de yüze ait dinamiklerden elde edilen davranışsal özniteliklerin birlikte kullanıldığı deneysel çalışmalar sonucunda ele edilen yüz tanımlama sonuçları Tablo 3.3 ve Tablo 3.4'te verilmiştir. Deneylerde en az iki videosu olan denekler kullanıldığı için, k kat çapraz doğrulama için k değeri 2 olarak seçilmiştir. Tablo 3.3 ve Tablo 3.4'teki ilk sütunda, test görüntülerine uygulanan bozulmalar için kullanılan parametreler parantez içinde verilmiştir. Üç DESA'nın yüz tanıma doğrulukları,



Şekil 3.5 Farklı görüntü bozulmalarına sahip UvA-NEMO veri tabanından alınan örnek yüz görüntülerinde tespit edilen yüz işaretleri gösterilmektedir. (Parametreler parantez içinde belirtilmiştir. Yukarıdan aşağıya ve soldan sağa: Toplamsal Aydınlatma (+125), Kontrast Değişimi [Aralık: 0.49-0.51], Gauss Bulanıklığı ($\sigma = 10$), Çarpımsal Aydınlatma ($\times$ 2.5), Gauss Gürültüsü (Var = 0.1), Tuz ve Biber Gürültüsü (Oran = 0.3) gösterilmektedir.)

tabloda VGGFace, VGGFace 2 ve ArcFace olarak etiketlenmiş olan sütun 2, 4 ve 6'da verilmiştir. Bu sütunlardaki sonuçlar incelendiği zaman, bozulma miktarı arttıkça yüz tanıma doğruluklarının önemli ölçüde düştüğünü görülebilmektedir. Sütun 3, 5 ve 7'de ise önerilen HYT sisteminin yüz tanıma doğrulukları verilmiştir. Hibrit Yüz Tanıma sistemleri, DESA'lardan elde edilen görünüme dayalı öznitelikler ile yüze ait dinamiklerden elde edilen davranışsal özniteliklerin birlikte kullanıldığı sistemleri ifade etmektedir. Önerilen HYT-VGGFace sisteminde, VGGFace'den elde edilen 4096 boyutlu öznitelik vektörü ile gülme videolarından elde edilen 128 boyutlu dinamik öznitelik vektörlerinin birlikte kullanılması ile 4224 boyutlu öznitelik vektörü elde edilmiştir. Önerilen HYT-VGGFace 2 sisteminde, VGGFace 2'den elde edilen 2048 boyutlu öznitelik vektörü ile gülme videolarından elde edilen 128 boyutlu dinamik öznitelik vektörlerinin birlikte kullanılması ile 2176 boyutlu öznitelik vektörü elde edilmiştir ve kişi tanıma için bu öznitelik vektörü kullanılmıştır. Sütun 7'de bulunan HYT-ArcFace sisteminde, ArcFace'den elde edilen 512 boyutlu öznitelik vektörü ile gülme videolarından elde edilen 128 boyutlu dinamik öznitelik vektörlerinin birlikte kullanılması ile 640 boyutlu öznitelik vektörü elde edilmiştir ve kişi tanıma için bu

öznitelik vektörü kullanılmıştır. Yüze ait dinamiklerden elde edilen özniteliklerin, yüz tanıma doğruluğunu iyileştirdiği durumlar kalın yazımla vurgulanmıştır.

Tablo 3.3 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için UvA-NEMO gülümseme veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GB" Gauss Bulanıklığı'nı, "TA" Toplamsal Aydınlatmayı, "ÇA" Çarpımsal Aydınlatmayı belirtmektedir.)

Görüntü Bozulmaları	**VGGFace**	**HYT-VGGFace**	**VGGFace 2**	**HYT-VGGFace 2**	**ArcFace**	**HYT-ArcFace**
Bozulma Yok	99.98	**100.00**	100.00	100.00	99.98	**100.00**
GB ($\sigma = 2.5$)	99.98	99.98	100.00	100.00	99.98	**100.00**
GB ($\sigma = 5$)	99.91±0.01	**99.93±0.01**	99.66±0.02	**99.82±0.02**	99.78±0.03	99.78±0.03
GB ($\sigma = 7.5$)	73.99±1.52	**79.16±0.38**	89.29±0.61	**92.72±0.09**	85.69±2.38	**91.72±0.20**
GB ($\sigma = 10$)	19.48±1.41	**24.13±0.41**	48.74±1.36	**56.07±0.21**	36.79±1.05	**50.90±0.23**
GB ($\sigma = 12.5$)	5.10±0.99	5.10±0.99	18.01±0.65	18.01±0.65	14.08±0.42	14.08±0.42
GB ($\sigma = 15$)	2.31±0.14	2.31±0.14	5.33±0.33	5.33±0.33	6.83±0.20	6.83±0.20
TA (5)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (10)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (15)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (20)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (25)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (30)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (50)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
TA (70)	99.91±0.01	**100.00**	100.00	100.00	99.98	**100.00**
TA (100)	99.57±0.02	**99.73±0.03**	99.91±0.01	**99.93±0.01**	99.98	**100.00**
TA (125)	97.21±0.08	**97.85±0.05**	99.55±0.05	**99.80±0.02**	99.91±0.01	**100.00**
TA (150)	82.40±0.57	82.40±0.57	95.19±0.06	95.19±0.06	98.30±0.05	98.30±0.05
TA (175)	37.33±3.33	37.33±3.33	63.96±0.16	63.96±0.16	78.04±0.21	78.04±0.21
TA (200)	5.40±0.40	5.40±0.40	11.86±0.16	11.86±0.16	28.01±0.89	28.01±0.89
ÇA (1)	99.98	**100.00**	100.00	100.00	99.98	**100.00**
ÇA (1.5)	99.77±0.03	**99.86±0.03**	99.98	99.98	99.98	**100.00**
ÇA (2)	97.57±0.03	**98.14±0.04**	99.48±0.04	**99.73±0.03**	99.98	**100.00**
ÇA (2.5)	84.99±0.29	**86.67±0.07**	96.17±0.17	**97.30±0.06**	98.44±0.04	**98.62±0.04**
ÇA (3)	60.83±2.16	60.83±2.16	82.97±1.39	82.97±1.39	91.11±0.72	91.11±0.72
ÇA (4)	25.54±2.28	25.54±2.28	46.36±0.66	46.36±0.66	63.05±1.05	63.05±1.05
ÇA (5)	9.82±1.95	9.82±1.95	20.89±0.73	20.89±0.73	39.74±0.47	39.74±0.47

Örneğin, Gauss bulanıklığı (σ =7,5) için VGGFace 2 ağının elde ettiği yüz tanıma doğruluğu %89,29'dur, ancak önerilen hibrit yöntem ile (HYT-VGGFace 2) yüz tanıma doğruluğu %92,72'lik bir başarıma ulaşmaktadır. Başka bir örnek, kontrast değişimi [0.49-0.51] için verilebilmektedir. VGGFace 2'nin yüz tanıma doğruluğu bu koşullar altında %77.66 iken, önerilen hibrit yöntem ile (HYT-VGGFace 2) %82.56'lık yüz tanıma doğruluğuna ulaşabilmektedir. Koyu renkle gösterilen durumlar incelendiği zaman, önerilen hibrit yöntem (HYT-VGGFace 2) için, (%0.05-%4.90) arasında değişen bir yüz tanıma doğruluk artışı sağladığı gözlemlenmiştir.

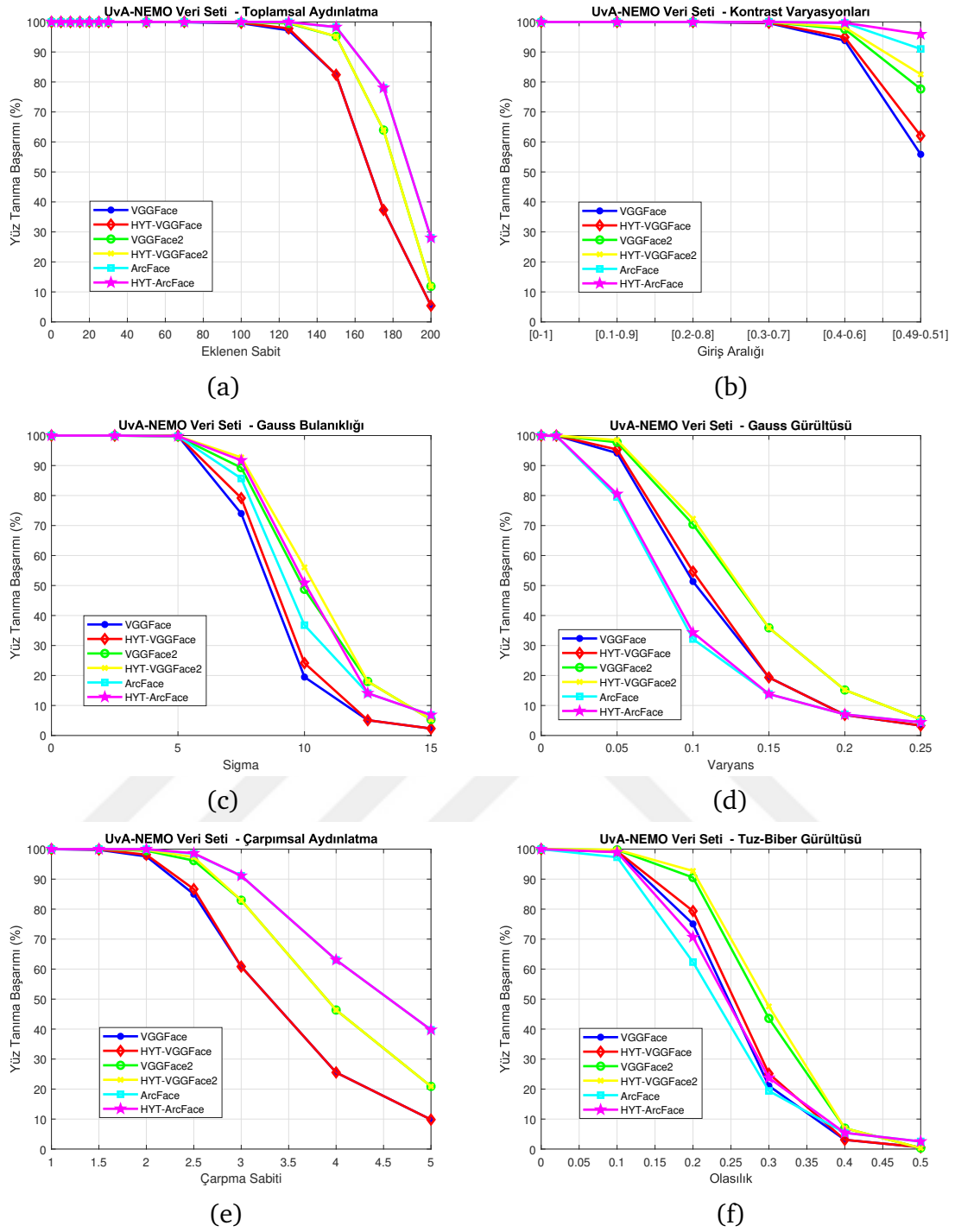
Şekil 3.6'de, karşılaştırma kolaylığı sağlamak amacıyla UvA-NEMO veri kümesi

Tablo 3.4 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için UvA-NEMO gülümseme veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GG" Gauss Gürültüsünü, "KD" Kontrast Değişimlerini ve "TBG" Tuz Biber Gürültüsünü belirtmektedir.)

Görüntü Bozulmaları	**VGGFace**	**HYT-VGGFace**	**VGGFace 2**	**HYT-VGGFace 2**	**ArcFace**	**HYT-ArcFace**
GG (var=0.01)	99.95±0.01	**99.98**	100.00	100.00	99.80±0.02	**99.91±0.01**
GG (var=0.05)	94.15±0.15	**95.35±0.05**	97.75±0.07	**98.50±0.05**	79.52±1.12	**80.52±0.42**
GG (var=0.1)	51.33±1.27	**54.66±1.06**	70.38±0.58	**72.26±0.48**	32.16±0.88	**34.32±0.62**
GG (var=0.15)	19.35±3.07	19.35±3.07	35.90±2.68	35.90±2.68	13.81±0.58	13.81±0.58
GG (var=0.2)	6.87±1.76	6.87±1.76	15.17±2.82	15.17±2.82	7.03±0.35	7.03±0.35
GG (var=0.25)	3.31±0.27	3.31±0.27	5.35±1.59	5.35±1.59	4.40±0.10	4.40±0.10
KD[0.1-0.9]	99.98	**100.00**	100.00	100.00	99.98	**100.00**
KD[0.2-0.8]	99.98	**100.00**	100.00	100.00	99.98	**100.00**
KD[0.3-0.7]	99.61±0.05	**99.73±0.03**	99.98	**100.00**	99.98	**100.00**
KD[0.4-0.6]	93.78±0.18	**94.92±0.06**	97.60±0.05	**98.16±0.06**	99.68±0.01	**99.73±0.03**
KD[0.49-0.51]	55.86±1.16	**62.05±0.55**	77.66±0.36	**82.56±0.16**	91.00±0.05	**95.92±0.12**
TBG (0.1)	99.09±0.06	**99.30±0.04**	99.75±0.03	**99.82±0.02**	97.30±0.30	**98.98±0.04**
TBG (0.2)	75.05±1.05	**79.36±0.36**	90.54±1.31	**92.72±0.19**	62.33±1.33	**70.65±0.65**
TBG (0.3)	21.80±1.80	**25.15±0.65**	43.57±0.57	**47.56±0.56**	19.39±1.39	**23.75±0.75**
TBG (0.4)	3.08±0.21	3.08±0.21	6.96±0.36	6.96±0.36	5.42±0.12	5.42±0.12
TBG (0.5)	0.60±0.06	0.60±0.06	0.34±0.04	0.34±0.04	2.52±0.07	2.52±0.07

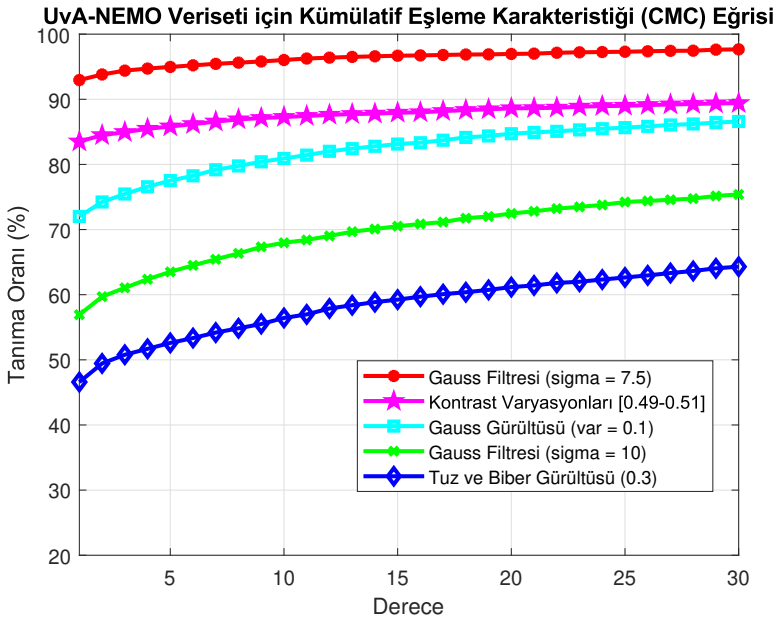
kullanılarak gerçekleştirilen deneysel çalışmarının sonuçlarının gösterildiği grafikler verilmiştir. Toplamsal aydınlatma, kontrast değişimleri ve çarpımsal aydınlatma değişimleri için ArcFace ile önerilen yöntemin (HYT-ArcFace) en iyi yüz tanıma başarımı elde ettiği görülmektedir. Öte yandan Gauss bulanıklığı, Gauss gürültüsü ve tuz-biber gürültüsü bozulmaları için ise, VGGFace 2 ile önerilen yöntem (HYT-VGGFace 2) en iyi yüz tanıma doğruluğunu vermektedir. Önerilen hibrit yüz tanıma sistemi kapalı küme tanımlama sistemi olduğundan, seçilen bozulmalar altında önerilen yöntemi değerlendirmek amacıyla HYT-VGGFace 2'nin kümülatif eşleşme karakteristik (CMC) eğrileri Şekil 3.7'da verilmiştir. CMC eğrileri, k'nin 1 ile 30 arasında değiştiği derecelere göre yüz tanımlama doğruluklarını göstermektedir. Eğri oluşturulurken kullanılan 5 farklı koşul, önerilen HYT sisteminin (HYT-VGGFace 2) DESA yöntemlerini kullanmaya kıyasla yüz tanıma doğruluğunda en yüksek gelişmeleri gösterdiği beş farklı koşuldur.

Deneysel çalışmalarda kullanılan bir diğer veri kümesi olan FEEDTUM veri kümesi



Şekil 3.6 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) ve HYT sistemler için UvA-NEMO gülümseme veri kümesini kullanılarak elde elde edilen yüz tanıma doğruluğu (%) grafikleri.

ise, 18 deneğin videolarını içeren nispeten küçük bir veri kümesidir. FEEDTUM veri kümesinden bir görüntü veri kümesi oluşturmak için her gülümseme videosundan (uzunluğundan bağımsız olarak) eşit aralıklarla 12 kare elde edilmiş ve böylelikle 54 videodan 648 adet görüntünün bulunduğu bir veri kümesi elde edilmiştir. Her deneğe ait bir videodan alınan görüntüler test için ayrılmış ve diğer iki videodan



Şekil 3.7 Beş farklı bozulma altında HYT-VGGFace 2 için UvA-NEMO gülümseme veri kümesini kullanılarak elde edilen Kümülatif Eşleşme Karakteristik (CMC) eğrisi.

elde edilen görüntüler eğitim için (kayıt verisi olarak) kullanılmıştır. Sonuç olarak, eğitim veri kümesi 432 adet görüntü içermekteyken, test veri kümesi 216 adet görüntü içermektedir. FEEDTUM veri kümesinde bulunan bazı gülümseme videolarının gülümseme ifadesinin bitiş segmentini içermediği, dolayısıyla nötr bir ifade ile bitmediğini gözlemlenmiştir. Ayrıca videolar incelendiği zaman bazı deneklerin duyguyu iyi ifade edemedikleri de gözlemlenmiştir. Bu nedenle, bu tür videolar için istatistiksel dinamik öznitelikleri çıkarmaktaki başarım, 3 duygu segmentini içeren videolarla kıyaslandığı zaman daha zordur.

FEEDTUM veri kümesindeki gülümseme videoları için iki yönlü çapraz doğrulama kullanılarak elde edilen yüz tanıma sonuçları, ilk sütunun parantez içinde belirtilen parametrelerle test görüntülerine uygulanan bozulmaların belirtildiği Tablo 3.5 ve Tablo 3.6'da gösterilmektedir. Sütun 3, 5 ve 7'de, önerilen HYT sisteminin yüz tanıma doğrulukları verilmiştir. Yüz dinamiği özniteliklerinin doğruluğu iyileştirdiği durumlar kalın yazımla vurgulanmıştır. Tablo 3.5 ve Tablo 3.6'daki sonuçlar incelendiği zaman, Yalnızca DESA'dan elde edilen öznitelikleri kullanarak yüz tanıma gerçekleştirildiği durumda, yüz tanıma başarımının önemli ölçüde düşebileceğini ve bazı durumlarda dinamik tabanlı özniteliklerin, olumsuz şartlardan oluşan performans kaybını telafi edebileceği görülmektedir. Örneğin, kontrast değişimleri [0.49-0.51] için, VGGFace 2 derin öğrenme ağı kullanılarak elde edilen yüz tanıma doğruluğu %74.07 iken, önerilen HYT-VGGFace 2 yöntemi ile yüz tanıma doğruluğunun %3.24 artışla %77.31 olarak elde edildiği görülmektedir.

FEEDTUM veri kümesi kullanılarak gerçekleştirilen deneysel çalışmaların sonuçlarının grafiksel gösterimi Şekil 3.8'de verilmiştir. Grafiklerde de görüldüğü üzere, bazı durumlarda, yüz dinamiklerini de kullanarak yüz tanıma gerçekleştiren önerilen hibrit yöntemin tek başına DESA'lardan elde edilen öznitelikleri kullanmaya kıyasla daha iyi sonuçlar verdiği gözlemlenebilmektedir. FEEDTUM veri kümesi, gülümseme ifadesinin başlangıç-tepe bölgesi-bitiş aşamalarını içermeyen videolar içermesi nedeniyle, hibrit yöntem ile yüz tanıma doğruluklarındaki artışlar UvA-NEMO veri tabanındaki kadar net değildir. Bu nedenle, gülümsemenin segmentlerinde bir eksiklik olması durumunda, dinamik öznitelikler de aynı derecede etkili olamayabileceği sonucu çıkarılabilmektedir.

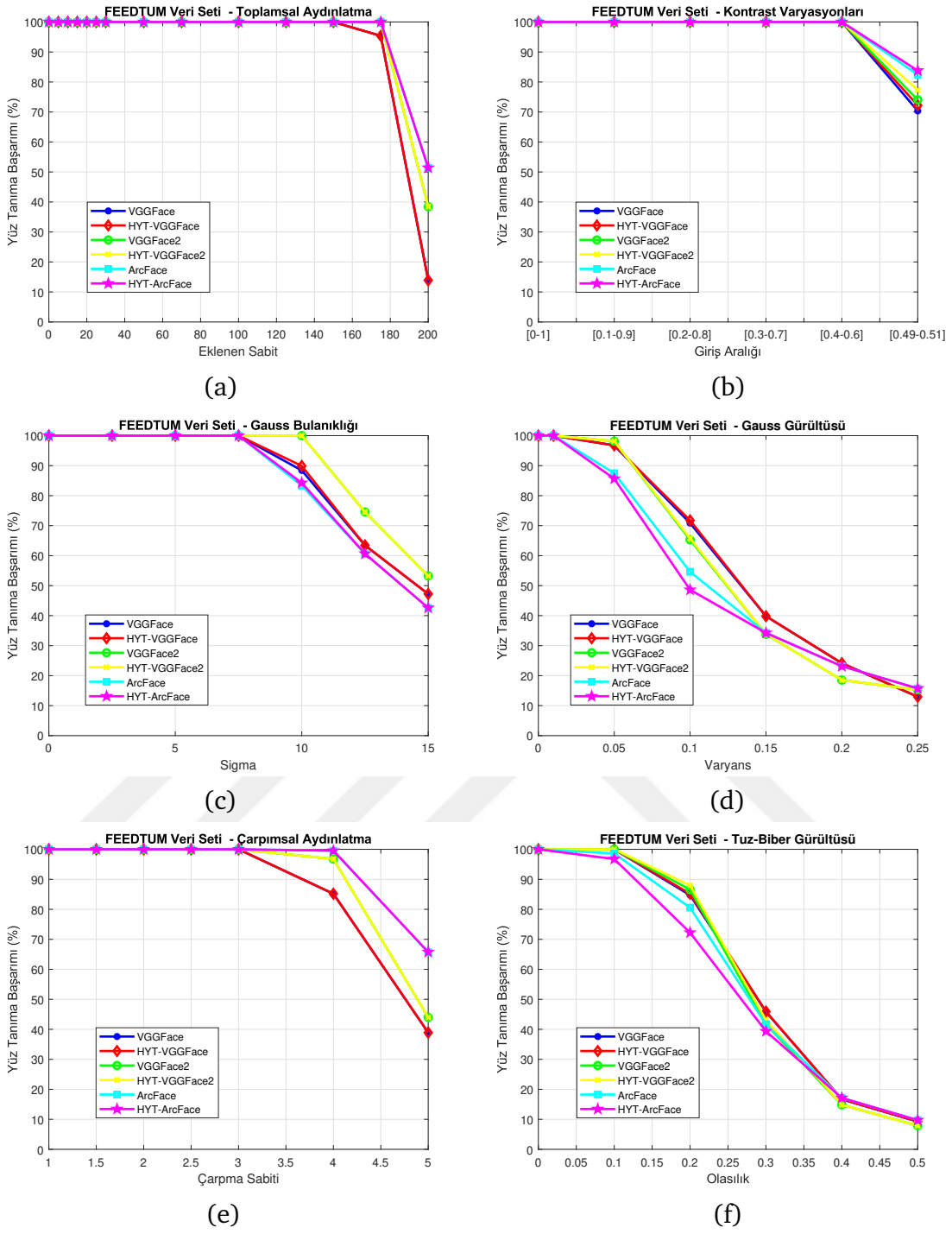
Önerilen hibrit yüz tanıma sistemi kapalı küme tanımlama sistemi olduğundan, seçilen bozulmalar altında önerilen yöntemi değerlendirmek amacıyla FEEDTUM veri kümesi kullanılarak kümülatif eşleşme karakteristik (CMC) eğrileri Şekil 3.9'de verilmiştir. Eğri oluşturulurken kullanılan 5 farklı koşul, önerilen HYT sisteminin (HYT-VGGFace) DESA yöntemlerini kullanmaya kıyasla yüz tanıma doğruluğunda en yüksek gelişmeleri gösterdiği beş farklı koşuldur.

Tablo 3.5 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için FEEDTUM veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GB" Gauss Bulanıklığı'nı, "TA" Toplamsal Aydınlatmayı, "ÇA" Çarpımsal Aydınlatmayı belirtmektedir.)

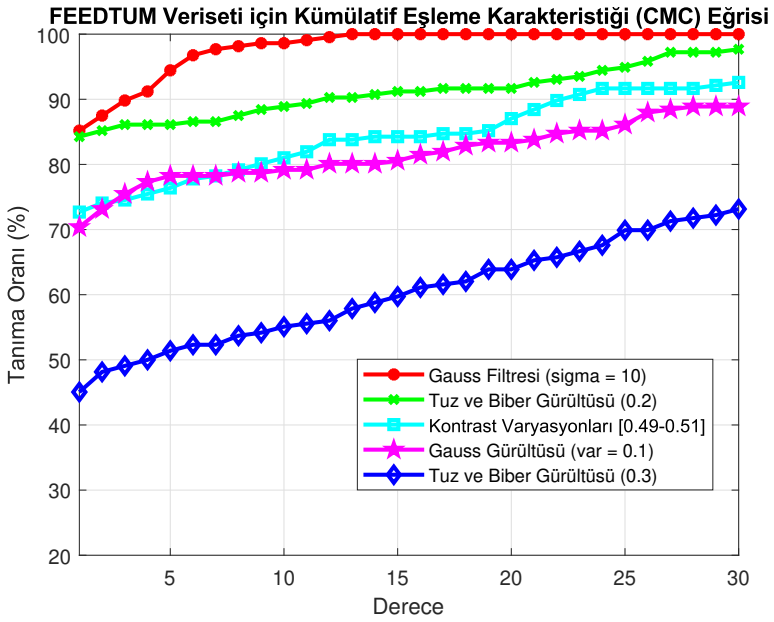
Görüntü Bozulmaları	**VGGFace**	**HYT-VGGFace**	**VGGFace 2**	**HYT-VGGFace 2**	**ArcFace**	**HYT-ArcFace**
Bozulma Yok	100.00	100.00	100.00	100.00	100.00	100.00
GB ($\sigma = 2.5$)	100.00	100.00	100.00	100.00	100.00	100.00
GB ($\sigma = 5$)	100.00	100.00	100.00	100.00	100.00	100.00
GB ($\sigma = 7.5$)	100.00	100.00	100.00	100.00	100.00	100.00
GB ($\sigma = 10$)	88.42±0.32	**89.81±0.19**	100.00	100.00	83.33±0.33	**84.26±0.26**
GB ($\sigma = 12.5$)	63.43±2.23	63.43±2.23	74.54±0.64	74.54±0.64	60.65±1.35	60.65±1.35
GB ($\sigma = 15$)	47.22±1.42	47.22±1.42	53.24±0.82	53.24±0.82	42.59±1.25	42.59±1.25
TA (5)	100.00	100.00	100.00	100.00	100.00	100.00
TA (10)	100.00	100.00	100.00	100.00	100.00	100.00
TA (15)	100.00	100.00	100.00	100.00	100.00	100.00
TA (20)	100.00	100.00	100.00	100.00	100.00	100.00
TA (25)	100.00	100.00	100.00	100.00	100.00	100.00
TA (30)	100.00	100.00	100.00	100.00	100.00	100.00
TA (50)	100.00	100.00	100.00	100.00	100.00	100.00
TA (70)	100.00	100.00	100.00	100.00	100.00	100.00
TA (100)	100.00	100.00	100.00	100.00	100.00	100.00
TA (125)	100.00	100.00	100.00	100.00	100.00	100.00
TA (150)	100.00	100.00	100.00	100.00	100.00	100.00
TA (175)	95.37±0.60	95.37±0.60	100.00	100.00	100.00	100.00
TA (200)	13.90±1.80	13.90±1.80	38.43±0.73	38.43±0.73	51.39±1.39	51.39±1.39
ÇA (1)	100.00	100.00	100.00	100.00	100.00	100.00
ÇA (1.5)	100.00	100.00	100.00	100.00	100.00	100.00
ÇA (2)	100.00	100.00	100.00	100.00	100.00	100.00
ÇA (2.5)	100.00	100.00	100.00	100.00	100.00	100.00
ÇA (3)	100.00	100.00	100.00	100.00	100.00	100.00
ÇA (4)	85.19±1.19	85.19±1.19	96.76±0.46	96.76±0.46	99.54±0.24	99.54±0.24
ÇA (5)	38.89±2.39	38.89±2.39	43.98±1.44	43.98±1.44	65.74±1.24	65.74±1.24

Tablo 3.6 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) için FEEDTUM veri kümesi kullanılarak elde edilen yüz tanıma doğruluğu (%). ("HYT" Hibrit Yüz Tanıma'yı, "GG" Gauss Gürültüsünü, "KD" Kontrast Değişimlerini ve "TBG" Tuz Biber Gürültüsünü belirtmektedir.)

Görüntü Bozulmaları	**VGGFace**	**HYT-VGGFace**	**VGGFace 2**	**HYT-VGGFace 2**	**ArcFace**	**HYT-ArcFace**
GG (var = 0.01)	100.00	100.00	100.00	100.00	100.00	100.00
GG (var = 0.05)	96.76±0.32	96.76±0.32	98.15±0.18	98.15±0.18	87.50±0.32	85.65±0.25
GG (var = 0.1)	70.84±0.82	**71.76±0.46**	65.28±0.38	**65.74±0.32**	54.63±0.47	48.61±0.29
GG (var = 0.15)	39.81±1.21	39.81±1.21	33.80±1.60	33.80±1.60	34.26±1.42	34.26±1.42
GG (var = 0.2)	24.15±2.52	24.15±2.52	18.52±0.90	18.52±0.90	23.15±2.60	23.15±2.60
GG (var = 0.25)	12.96±1.42	12.96±1.42	15.28±0.64	15.28±0.64	15.74±0.76	15.74±0.76
KD [0.1-0.9]	100.00	100.00	100.00	100.00	100.00	100.00
KD [0.2-0.8]	100.00	100.00	100.00	100.00	100.00	100.00
KD [0.3-0.7]	100.00	100.00	100.00	100.00	100.00	100.00
KD [0.4-0.6]	100.00	100.00	100.00	100.00	100.00	100.00
KD [0.49-0.51]	70.37±0.42	**72.22±0.32**	74.07±0.47	**77.31±0.46**	82.41±0.56	**83.80±0.30**
TBG (0.1)	100.00	100.00	100.00	100.00	98.61±0.09	96.76±0.23
TBG (0.2)	84.72±0.44	**85.19±0.21**	86.57±0.47	**87.96±0.34**	80.56±0.74	72.22±1.22
TBG (0.3)	45.83±1.21	**45.91±1.26**	41.67±2.40	**43.06±1.70**	41.67±2.34	39.35±1.24
TBG (0.4)	16.67±2.37	16.67±2.37	14.81±1.10	14.81±1.10	17.13±0.73	17.13±0.73
TBG (0.5)	9.26±1.66	9.26±1.66	7.87±2.10	7.87±2.10	9.72±1.10	9.72±1.10



Şekil 3.8 Çeşitli görüntü bozulmaları altında karşılaştırılan üç DESA (VGGFace, VGGFace 2 ve ArcFace) ve HYT sistemler için FEEDTUM veri kümesini kullanılarak elde elde edilen yüz tanıma doğruluğu (%) grafikleri.



Şekil 3.9 Beş farklı bozulma altında HYT-VGGFace 2 için FEEDTUM veri kümesini kullanılarak elde edilen Kümülatif Eşleşme Karakteristik (CMC) eğrisi.

4
SONUÇ VE ÖNERİLER

Bu tez kapsamında, literatürde yumuşak biyometri olarak adlandırılan davranışsal özelliklerin elde edilmesi ve yüz tanıma performanslarına etkilerinin incelenmesi üzerine bir çalışma gerçekleştirilmiştir. Bu kapsamda, öncelikle literatürde yüz tanıma sistemleri hakkında detaylı bir araştırma gerçekleştirilmiş, son 10 yıl içerisinde literatürde yüz tanıma için kullanılan yöntemler ve test sonuçları incelenmiştir. Bu inceleme sonucunda, derin öğrenme ağlarının kullanıldığı yüz tanıma sistemlerinin, özellikle 2010'lu yıllardan itibaren yüz tanıma performansının insan yüz tanıma performansı yaklaştığı, hatta 2018'li yıllardan sonra insandan daha yüksek performans gösterdiği anlaşılmıştır. LFW veri tabanı kullanılarak gerçekleştirilen yüz tanıma test performanslarına bakıldığında, 2014 yılı ile 2019 yılları arasında yüz tanıma sistemlerinin performansında olağanüstü bir artış olduğu (%97,35'ten %99,86'ya) gözlemlenmektedir.

Bununla birlikte, son dönemlerde gerçekleştirilen çalışmalar [14, 198], farklı aydınlatma, baş açısı, gürültü ve düşük çözünürlük gibi zorlu koşulların derin öğrenme ağlarının yüz tanıma performansı üzerinde olumsuz etkileri olduğunu göstermiştir. Bu problemin üstesinden gelmek amacıyla, bazı araştırmacılar görüntünün derin sinir ağına verilmeden önce ön işlemeye tabi tutulması üzerine yoğunlaşırken, bazı araştırmacılar ise davranışsal biyometrinin yardımcı olarak kullanılabilmesine yoğunlaşmışlardır. Bu tez kapsamında da, bu problemi çözmek amacıyla, derin öğrenme ağlarından elde edilen görünüm-temelli öznitelik vektörlerinin, gülümseme yüz ifadelerinden elde edilen davranışsal-temelli öznitelik vektörleri ile birlikte kullanıldığı yeni bir yöntem önerilmiştir. Önerilen hibrit yöntemin başarımını incelemek amacıyla, gülümseme videolarının bulunduğu iki farklı veri kümesi kullanılarak deneysel çalışmalar gerçekleştirilmiştir. Gerçekleştirilen deneysel çalışmalar sonucunda, gülümseme yüz ifadesinden elde edilen istatistiksel dinamik özniteliklerin, zorlu koşullar altında yüz tanıma başarımı azalan derin sinir ağlarına yardımcı olarak kullanılması durumunda, yüz tanıma performansının artırdığı gözlemlenmiştir. Bununla birlikte, istatistiksel dinamik özniteliklerin başarılı bir

şekilde elde edilebilmesi için duygunun bölütlenmesinin önemi de deneysel çalışmalar sonucunda ortaya çıkmıştır.

Bu tez çalışması kapsamında,

- Literatürde görüntü ve video veri kümeleri ve değerlendirme ölçütleri de dahil olmak üzere görüntü ve video tabanlı yüz tanıma üzerine yapılan çok sayıda çalışmaya ilişkin güncel, kapsamlı ve kompakt bir literatür taraması sunulmuştur.
- Seyreklik ve derin öğrenme tabanlı yöntemler gibi yeni yöntemleri de içeren görüntü ve video tabanlı yöntemlerin bulunduğu genel bir taksonomi oluşturulmuştur. Taksonomi oluşturmanın amacı, literatürde yüz tanıma için kullanılan yöntemlere genel bir bakış sağlamaktır.
- Büyük ölçekli ve kısıtlanmamış zorlu veri setleri kullanılarak gerçekleştirilen yüz tanıma çalışmalarında dikkate değer performanslar gösteren derin öğrenmeye dayalı yöntemler ile ilgili incelemeler gerçekleştirilmiştir. Bu sayede, 2010'lu yılların başından itibaren, yüz tanımaya yeni bir bakış açısı getiren derin öğrenme temelli yöntemler hakkında detaylı bilgi verilmiştir.
- Son zamanlarda önerilmiş DESA tabanlı üç yüz tanıma yönteminin (VGGFace, VGGFace2 ve ArcFace) çeşitli parametrelerle simüle edilmiş altı farklı olumsuz koşul altında kapsamlı bir performans karşılaştırması verilmiştir.
- Görünüme dayalı özniteliklerle birlikte istatistiksel dinamik özniteliklerin kullanılmasının, olumsuz koşullar altında (ciddi görüntü bozulmalarında) yüz tanımanın gürbüzlüğünü geliştirdiği iki farklı gülümseme video veri kümeleri kullanılarak gerçekleştirilen deneysel çalışmalar ile gösterilmiştir.

Yüz tanıma sistemlerinin geliştirilmesi ile ilgili gelecekteki çalışmalarda, aşağıdaki kavramların ön planda olacağı öngörülmektedir:

Görüntü iyileştirme: Görüntü iyileştirme çalışmalarındaki amaç, yüz tanıma sistemlerinin performansını artırmak için düşük çözünürlüklü yüz görüntülerine süper çözünürlüklü algoritmaları veya 3 boyutlu görüntü oluşturma algoritmalarını uygulamaktır. Güvenlik kameralarından elde edilen büyük miktarda veri olduğundan, düşük çözünürlüklü görüntülerin kullanıldığı yüz tanıma, üzerinde çalışılması önem arz eden ve zorlu bir araştırma problemidir. Düşük çözünürlüklü ve yüksek çözünürlüklü görüntüler arasındaki boşluğu azaltmak için çözünürlük temelli yeni öznitelik çıkarma yöntemlerine ihtiyaç vardır.

Kayıp fonksiyonları: Derin sinir ağlarını kullanan yüz tanıma sistemlerinin olumsuz koşullar altında performansları düştüğünden, bir başka araştırma alanı ise yeni kayıp fonksiyonları kullanılarak bu sistemlerin performansının arttırılmasıdır. Halihazırda, derin öğrenme tabanlı yüz tanıma ve yüz yanıltma saldırıları önleme sistemleri için literatürde kullanılan yaklaşık 20 farklı kayıp fonksiyonu bulunmaktadır. Gelecekte yüz tanıma sistemlerinde kullanılan kayıp fonksiyonlarının sayısının artacağı öngörülmektedir.

Veri kümesi tasarımı: Derin sinir ağlarını kullanarak yüz tanımanın gürbüzlüğünü geliştirmek için eğitim sırasında farklı aydınlatma, poz ve gürültü efektlerine sahip görüntüler kullanılabilir. Açıklama içeren büyük veri kümelerini elde etmek çok zor olduğu için, aktif öğrenme ve benzeri yaklaşımları kullanarak eğitim görüntülerinin sayısını azaltmak başka bir yaklaşım olarak ortaya çıkmaktadır. Bir başka araştırma konusunda olarak da, sadece yüz ifadelerinden elde edilen dinamik öznitelikler kullanılarak kişinin kimlik bilgisini elde etmek için aynı kişinin çeşitli yüz ifadelerini içeren yeni çok kipli video tabanlı veri kümelerine de ihtiyaç vardır.

Yumuşak biyometri: Yumuşak biyometrik veriler, yüz videolarındaki uzamsal-zamansal bilgiler kullanılarak yüz dinamiklerinden çıkarılmaktadır. Araştırma sonuçları, yumuşak biyometrik özniteliklerin tek başına yüksek başarımlı yüz tanıma için yeterli olmadığını gösterse de, olumsuz koşullar altında sistemin yüz tanıma doğruluğunu artırmak için görünüm tabanlı yöntemlerle birlikte kullanılabilirler. Duygusal ifadelerin (öfke, üzüntü, şaşkınlık, iğrenme, korku gibi) yüz dinamiklerinin kimlik bilgisi taşıyıp taşımadığı hala araştırmacıların ilgisini çeken bir konudur.

Yüz yanıltma saldırıları önleme: Derin sinir ağlarını kullanan yüz tanıma sistemleri, çeşitli senaryolarda insan performansını aştığını gösterse de, derin öğrenme ağlarının insanlardan daha kolay yanıltıldığı gözlemlenmiştir. Bu nedenle, yüz tanımayı yanıltma saldırılarına karşı daha sağlam hale getirmek güncel bir araştırma yönüdür. Daha gürbüz derin sinir ağların geliştirilmesine yönelik araştırmaları teşvik etmek için belirli zamanlarda çeşitli yarışmalar düzenlenmektedir. Son zamanlarda, yüz yanıltma saldırılarını önleme için başarılı performans gösteren birçok derin sinir ağı tabanlı yöntem önerilmiştir. Gelecekte, yeni yüz yanıltma saldırıları türleri oluşabileceğinden, sıfır atışlı yüz yanıltma saldırıları önleme yaklaşımlarına ihtiyaç duyulacağı beklenmektedir. Derin sinir ağlarının düşmanca ve anlamsal saldırılara karşı gürbüzlüğünün geliştirilmesi de aktif bir araştırma alanıdır.

Çok kipli ve çapraz kipli yüz tanıma: Bu tezde, yüz bilgilerini yakalamak için en yaygın kullanılan ve uygun maliyetli yol olduğu için görsel (RGB) modalitesini

kullanarak görüntü ve video tabanlı yüz tanımaya odaklandık. Bununla birlikte, diğer modaliteleri (3B, yakın kızılötesi, termal kızılötesi, eskizler) kullanarak yüz tanıma ve çok kipli yüz tanıma ve yüz yanıltma saldırılarını önlemenin aktif ve ilginç araştırma alanları olduğu belirtilmelidir. Farklı kipler kullanılarak elde edilen yüz görüntülerini eşleştirmeye çalışan heterojen yüz tanıma, araştırmacıların ilgisini çekmektedir. Örneğin yüz eskizlerini fotoğraflarla eşleştirmek adli güvenlik için önemli bir sorundur. Bir eskizden bir yüz fotoğrafını sentezlemek ve bunun tersini yapmak, araştırmak için ilginç ve zorlu problemlerdir.

Yukarıdaki kavramları içeren literatürdeki akademik çalışmaların nihai hedefi, insan görme sisteminin başarımını aşabilen otomatik bir yüz tanıma sistemi geliştirmektir. Bu amaca, bilgisayarlı görme araştırmacıları ve sinirbilimciler arasındaki karşılıklı ve koordineli çalışmaları ile ulaşılabilir.

KAYNAKÇA

[1] A. Dantcheva, P. Elia, A. Ross, "What else does your biometric data reveal? a survey on soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 441–467, 2015.

[2] H. Dibeklioğlu, F. Alnajar, A. A. Salah, T. Gevers, "Combining facial dynamics with appearance for age estimation," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1928–1943, 2015.

[3] A. Dantcheva, F. Brémond, "Gender estimation based on smile-dynamics," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 719–729, 2016.

[4] C. E. Erdem, C. Turan, Z. Aydin, "Baum-2: A multilingual audio-visual affective face database," *Multimedia tools and applications*, vol. 74, no. 18, pp. 7429–7459, 2015.

[5] S. Zhalehpour, Z. Akhtar, C. E. Erdem, "Multimodal emotion recognition based on peak frame selection from video," *Signal, Image and Video Processing*, vol. 10, no. 5, pp. 827–834, 2016.

[6] B. Martinez, M. F. Valstar, B. Jiang, M. Pantic, "Automatic analysis of facial actions: A survey," *IEEE transactions on affective computing*, vol. 10, no. 3, pp. 325–347, 2017.

[7] S. Zhalehpour, O. Onder, Z. Akhtar, C. E. Erdem, "Baum-1: A spontaneous audio-visual face database of affective and mental states," *IEEE Transactions on Affective Computing*, vol. 8, no. 3, pp. 300–313, 2016.

[8] L. M. Vaina, J. Solomon, S. Chowdhury, P. Sinha, J. W. Belliveau, "Functional neuroanatomy of biological motion perception in humans," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 11 656–11 661, 2001.

[9] V. Bruce, P. J. Hancock, A. M. Burton, "Human face perception and identification," *Face recognition*, pp. 51–72, 1998.

[10] V. Bruce, A. Young, "Understanding face recognition," *British journal of psychology*, vol. 77, no. 3, pp. 305–327, 1986.

[11] G. Davies, H. Ellis, J. Shepherd, "Cue saliency in faces as assessed by the 'photofit'technique," *Perception*, vol. 6, no. 3, pp. 263–269, 1977.

[12] H. Hill, V. Bruce, "Effects of lighting on matching facial surfaces,"," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22, pp. 986–1004, 1996.

[13] B. Knight, A. Johnston, "The role of movement in face recognition," *Visual cognition*, vol. 4, no. 3, pp. 265–273, 1997.

[14] K. Grm, V. Štruc, A. Artiges, M. Caron, H. K. Ekenel, "Strengths and weaknesses of deep learning models for face recognition against image degradations," *Iet Biometrics*, vol. 7, no. 1, pp. 81–89, 2018.

[15] G. Pala, C. E. Erdem, "Performance comparison of deep learning based face identification methods for video under adverse conditions," in *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, IEEE, 2019, pp. 90–97.

[16] J. Mohapatra, T.-W. Weng, P.-Y. Chen, S. Liu, L. Daniel, "Towards verifying robustness of neural networks against a family of semantic perturbations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 244–252.

[17] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, "Face recognition: A literature survey," *ACM computing surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[18] T. Kanade *et al.*, *Computer recognition of human faces*. Birkhäuser Basel, 1977, vol. 47.

[19] M. Wang, W. Deng, "Deep face recognition: A survey," *arXiv:1804.06655*, 2019.

[20] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[21] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2008.

[22] M. A. Turk, A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition*, IEEE Computer Society, 1991, pp. 586–587.

[23] W. Deng, J. Hu, J. Lu, J. Guo, "Transform-invariant pca: A unified approach to fully automatic facealignment, representation, and recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 6, pp. 1275–1284, 2013.

[24] X. He, S. Yan, Y. Hu, P. Niyogi, H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[25] J. Yang, D. Zhang, J.-y. Yang, B. Niu, "Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 650–664, 2007.

[26] W. Deng, J. Hu, J. Guo, "Extended src: Undersampled face recognition via intraclass variant dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864–1870, 2012.

[27] W. Deng, J. Hu, J. Guo, "Face recognition via collaborative representation, ıts discriminant nature and superposed representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 10, pp. 2513–2521, 2017.

[28] C. Liu, H. Wechsler, "Independent component analysis of gabor features for face recognition," *IEEE transactions on Neural Networks*, vol. 14, no. 4, pp. 919–928, 2003.

[29] L. Xue-fang, P. Tao, "Realization of face recognition system based on gabor wavelet and elastic bunch graph matching," in *2013 25th Chinese Control and Decision Conference (CCDC)*, IEEE, 2013, pp. 3384–3386.

[30] K. Meena, A. Suruliandi, "Local binary patterns and its variants for face recognition," in *2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, IEEE, 2011, pp. 782–786.

[31] T. Ahonen, A. Hadid, M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[32] W. Deng, J. Hu, J. Guo, "Compressive binary patterns: Designing a robust binary face descriptor with random-field eigenfilters," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 3, pp. 758–767, 2018.

[33] Z. Cao, Q. Yin, X. Tang, J. Sun, "Face recognition with learning-based descriptor," in *2010 IEEE Computer society conference on computer vision and pattern recognition*, IEEE, 2010, pp. 2707–2714.

[34] Z. Lei, M. Pietikäinen, S. Z. Li, "Learning discriminant face descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 289–302, 2013.

[35] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[36] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701–1708.

[37] O. M. Parkhi, A. Vedaldi, A. Zisserman, "Deep face recognition," in *BMVC*, 2015. [Online]. Available: `https://www.robots.ox.ac.uk/~vgg/data/vgg_face/`.

[38] R. Chellappa, C. L. Wilson, S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, 1995.

[39] A. Samal, P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern recognition*, vol. 25, no. 1, pp. 65–77, 1992.

[40] A. F. Abate, M. Nappi, D. Riccio, G. Sabatino, "2d and 3d face recognition: A survey," *Pattern recognition letters*, vol. 28, no. 14, pp. 1885–1906, 2007.

[41] R. Jafri, H. R. Arabnia, "A survey of face recognition techniques," *journal of information processing systems*, vol. 5, no. 2, pp. 41–68, 2009.

[42] M. Hassaballah, S. Aly, "Face recognition: Challenges, achievements and future directions," *IET Computer Vision*, vol. 9, no. 4, pp. 614–626, 2015.

[43] Z. Mahmood, N. Muhammad, N. Bibi, T. Ali, "A review on state-of-the-art face recognition approaches," *Fractals*, vol. 25, no. 02, p. 1 750 025, 2017.

[44] X. Zhang, Y. Gao, "Face recognition across pose: A review," *Pattern Recognition*, vol. 42, no. 11, pp. 2876–2896, 2009.

[45] C. Ding, D. Tao, "A comprehensive survey on pose-ınvariant face recognition," *ACM Transactions on Intelligent Systems and Technology*, vol. 7, no. 3, 2016.

[46] X. Zou, J. Kittler, K. Messer, "Illumination invariant face recognition: A survey," in *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 2007.

[47] M. Tistarelli, M. Bicego, E. Grosso, "Dynamic face recognition: From human to machine vision," *Image and Vision Computing*, vol. 27, no. 3, pp. 222–232, 2009. DOI: DOI:10.1016/j.imavis.2007.05.006.

[48] H. Wang, Y. Wang, Y. Cao, "Video-based face recognition: A survey," *International Journal of Computer and Information Engineering*, vol. 3, no. 12, pp. 293–302, 2009.

[49] F. Matta, J. Dugelay, "Person recognition using facial video information: A state of the art," *Journal of Visual Languages and Computing*, vol. 20, no. 3, pp. 180–187, 2009. DOI: DOI:10.1016/j.jvlc.2009.01.002.

[50] A. Hadid, J. Dugelay, M. Pietikainen, "On the use of dynamic features in face biometrics: Recent advances and challenges," *Signal, Image and Video Processing*, vol. 5, pp. 495–506, 2011. DOI: DOI:10.1007/s11760-011-0247-3.

[51] J. R. Barr, K. W. Bowyer, P. J. Flynn, S. Biswas, "Face recognition from video: A review," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 26, no. 5, pp. 1–53, 2012. DOI: DOI:10.1142/S0218001412660024.

[52] A. Scheenstra, A. Ruifrok, R. C. Veltkamp, "A survey of 3d face recognition methods," in *International Conference on Audio and Video based Biometric Person Authentication*, 2005, pp. 891–899.

[53] K. W. Bowyer, K. Chang, P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition," *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[54] H. Zhou, A. Mian, L. Wei, D. Creighton, M. Hossny, S. Nahavandi, "Recent advances on singlemodal and multimodal face recognition : A survey," *EEE Trans. on Human-Machine Systems*, vol. 44, no. 6, pp. 701–716, 2014.

[55] R. Ramachandra, C. Busch, "Presentation attack detection methods for face recognition systems," *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–37, 2017.

[56] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, C. Busch, "Face recognition systems under morphing attacks: A survey," *IEEE Access*, vol. 7, pp. 23 012–23 026, 2019, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2899367.

[57] Y. Xu, Z. Li, J. Yang, D. Zhang, "A survey of dictionary learning algorithms for face recognition," *IEEE Access*, vol. 5, pp. 8502–8514, 2017.

[58] S. Balaban, "Deep learning and face recognition: The state of the art," in *Proceedings, Biometric and Surveillance Technology for Human and Activity*, vol. 9457, 2015.

[59] I. Masi, Y. Wu, T. Hassner, P. Natarajan, "Deep face recognition: A survey," in *SIBGRAPI - Conference on Graphics, Patterns and Images*, 2018.

[60] R. Ranjan, S. Sankaranarayanan, A. Bansal, N. Bodla, J. C. Chen, V. M. Patel, C. D. Castillo, R. Chellappa, "Deep learning for understanding faces: Machines may be just as good, or better, than humans," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 66–83, 2018.

[61] A. K. Jain, A. A. Ross, K. Nandakumar, *Introduction to biometrics*. Springer Science & Business Media, 2011.

[62] H. Dibeklioğlu, A. A. Salah, T. Gevers, "Are you really smiling at me? spontaneous versus posed enjoyment smiles," in *European Conference on Computer Vision*, Springer, 2012, pp. 525–538.

[63] R. Ranjan, S. Sankaranarayanan, A. Bansal, N. Bodla, J.-C. Chen, V. M. Patel, C. D. Castillo, R. Chellappa, "Deep learning for understanding faces: Machines may be just as good, or better, than humans," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 66–83, 2018.

[64] P. Viola, M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.

[65] W. Chen, K. Wang, H. Jiang, M. Li, "Skin color modeling for face detection and segmentation: A review and a new approach," *Multimedia Tools and Applications*, vol. 75, no. 2, pp. 839–862, 2016.

[66] Y. Tayal, R. Lamba, S. Padhee, "Automatic face detection using color based segmentation," *International Journal of Scientific and Research Publications*, vol. 2, no. 6, pp. 1–7, 2012.

[67] C. Erdem, S. Ulukaya, A. Karaali, A. T. Erdem, "Combining haar feature and skin color based classifiers for face detection," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2011, pp. 1497–1500.

[68] Y. Li, B. Sun, T. Wu, Y. Wang, "Face detection with end-to-end integration of a convnet and a 3d model," in *European Conference on Computer Vision*, Springer, 2016, pp. 420–436.

[69] P. Hu, D. Ramanan, "Finding tiny faces," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 951–959.

[70] H. Jiang, E. Learned-Miller, "Face detection with the faster r-cnn," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, IEEE, 2017, pp. 650–657.

[71] S. Ren, K. He, R. Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv:1506.01497*, 2015.

[72] S. S. Farfade, M. J. Saberian, L.-J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 2015, pp. 643–650.

[73] S. Yang, P. Luo, C.-C. Loy, X. Tang, "From facial parts responses to face detection: A deep learning approach," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 3676–3684.

[74] H. Li, Z. Lin, X. Shen, J. Brandt, G. Hua, "A convolutional neural network cascade for face detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 5325–5334.

[75] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, Springer, 2016, pp. 21–37.

[76] S. Yang, Y. Xiong, C. C. Loy, X. Tang, "Face detection through scale-friendly deep convolutional networks," *arXiv preprint arXiv:1706.02863*, 2017.

[77] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, S. Z. Li, "S3fd: Single shot scale-invariant face detector," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 192–201.

[78] A. Bansal, C. Castillo, R. Ranjan, R. Chellappa, "The do's and don'ts for cnn-based face verification," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 2545–2554.

[79] V. Kazemi, J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1867–1874.

[80] X. Xiong, F. De la Torre, "Supervised descent method and its applications to face alignment," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 532–539.

[81] J. Zhang, S. Shan, M. Kan, X. Chen, "Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment," in *European conference on computer vision*, Springer, 2014, pp. 1–16.

[82] S. Ren, X. Cao, Y. Wei, J. Sun, "Face alignment at 3000 fps via regressing local binary features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1685–1692.

[83] M. Bodini, "A review of facial landmark extraction in 2d images and videos using deep learning," *Big data and cognitive computing*, vol. 3, no. 1, p. 14, 2019.

[84] O. Çeliktutan, S. Ulukaya, B. Sankur, "A comparative study of face landmarking techniques," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, pp. 1–27, 2013.

[85] G. G. Chrysos, E. Antonakos, P. Snape, A. Asthana, S. Zafeiriou, "A comprehensive performance evaluation of deformable face tracking "in-the-wild"," *International Journal of Computer Vision*, vol. 126, no. 2, pp. 198–232, 2018.

[86] X. Jin, X. Tan, "Face alignment in-the-wild: A survey," *Computer Vision and Image Understanding*, vol. 162, pp. 1–22, 2017.

[87] D. Rathod, A. Vinay, S. Shylaja, S. Natarajan, "Facial landmark localization-a literature survey," *Int. J. Current Eng. Technol.*, vol. 4, no. 3, pp. 1901–1907, 2014.

[88] E. N. Sandıkcı, Ç. E. Erdem, S. Ulukaya, "A comparison of facial landmark detection methods," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2018, pp. 1–4.

[89] N. Wang, X. Gao, D. Tao, H. Yang, X. Li, "Facial feature point detection: A comprehensive survey," *Neurocomputing*, vol. 275, pp. 50–65, 2018.

[90] Y. Wu, Q. Ji, "Facial landmark detection: A literature survey," *International Journal of Computer Vision*, vol. 127, no. 2, pp. 115–142, 2019.

[91] Y. Wu, T. Hassner, K. Kim, G. Medioni, P. Natarajan, "Facial landmark detection with tweaked convolutional neural networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 12, pp. 3067–3074, 2017.

[92] J. Yang, Q. Liu, K. Zhang, "Stacked hourglass network for robust facial landmark localisation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 79–87.

[93] R. Ranjan, V. M. Patel, R. Chellappa, "Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 1, pp. 121–135, 2017.

[94] R. Ranjan, S. Sankaranarayanan, C. D. Castillo, R. Chellappa, "An all-in-one convolutional neural network for face analysis," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, IEEE, 2017, pp. 17–24.

[95] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[96] Y. Lin, S. Cheng, J. Shen, M. Pantic, "Mobiface: A novel dataset for mobile face tracking in the wild," in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, IEEE, 2019, pp. 1–8.

[97] F. S. Samaria, A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *IEEE Workshop on Applications of Computer Vision*, Dec. 1994, pp. 138–142. DOI: 10.1109/ACV.1994.341300. [Online]. Available: http://cam-orl.co.uk/facedatabase.html.

[98] M. J. Black, Y. Yacoob, "Recognizing facial expressions in image sequences using local parameterized models of image motion," *International Journal of Computer Vision*, vol. 25, no. 1, pp. 23–48, 1997.

[99] P. W. Hallinan, "A deformable model for the recognition of human faces under arbitrary ıllumination," Ph.D. dissertation, Harvard University, Cambridge, MA, USA, 1995. [Online]. Available: ftp://ftp.hrl.harvard.edu/pub/faces.

[100] P. J. Phillips, H.Wechsler, J. Huang, P. Rauss, "The feret database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing*, vol. 16, no. 5, pp. 295–306, 1998. [Online]. Available: https://nist.gov/itl/products-and-services/color-feret-database.

[101] A. Georghiades, P. Belhumeur, D. Kriegman, *Yale face database*, [Online] http://cvc.yale.edu/projects/yalefaces/yalefa, Center for computational Vision and Control at Yale University, 1997.

[102] M. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, "Coding facial expressions with gabor wavelets," in *IEEE international Conference on Automatic Face and Gesture Recognition*, 1998, pp. 200–205.

[103] D. B. Graham, N. M. Allinson, "Characterizing virtual eigensignatures for general purpose face recognition," in *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. Fogelman-Soulie, T. S. Huang, Eds., vol. 163, NATO ASI Series F, Computer and Systems Sciences: The name of the publisher, 1998, pp. 446–456. [Online]. Available: http://images.ee.umist.ac.uk/danny/database.html.

[104] A. Martinez, R. Benavente, "The ar face database," CVC Technical Report, Tech. Rep. 24, Jun. 1998. [Online]. Available: http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html.

[105] A. Nefian, M. H. Hayes III, "A hidden markov model-based approach for face detection and recognition," Ph.D. dissertation, School of Electrical and Computer Engineering, Georgia Institute of Technology, 1999. [Online]. Available: http://www.anefian.com/research/face_reco.htm.

[106] M. Soriano, E. Marszalec, M. Pietikainen, "Physics-based face database for color research," *Journal of Electronic Imaging*, vol. 9, no. 1, pp. 32–38, 2000. [Online]. Available: http://www.cse.oulu.fi/CMV/Downloads/Pbfd.

[107] T. Sim, S. Baker, M. Bsat, "The cmu pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2003.1251154. [Online]. Available: http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html.

[108] O. Jesorsky, K. J. Kirchberg, R. W. Frischholz, "Robust face detection using the hausdorff distance," in *Audio- and Video-Based Biometric Person Authentication*, J. Bigun, F. Smeraldi, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 90–95, ISBN: 978-3-540-45344-4. [Online]. Available: http://www.humanscan.de/support/downloads/facedb.php.

[109] D. A. Socolinsky, L. B. Wolff, J. D. Neuheisel, C. K. Eveland, "Illumination invariant face recognition using thermal infrared imagery," in *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, vol. 1, 2001, pp. I–I.

[110] A. S. Georghiades, P. N. Belhumeur, D. J. Kriegman, "From few to many: İllumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, Jun. 2001, ISSN: 0162-8828. DOI: 10.1109/34.927464. [Online]. Available: http://vision.ucsd.edu/content/extended-yale-face-database-b-b.

[111] P. J. Phillips, "Human identification technical challenges," in *Proceedings. International Conference on Image Processing*, vol. 1, Sep. 2002, pp. I–I. DOI: 10.1109/ICIP.2002.1037956. [Online]. Available: http://www.nd.edu/cvrl/HID-data.html.

[112] V. Jain, A. Mukherjee, *The ındian face database*, 2002. [Online]. Available: http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/.

[113] H. Bon-Woo, H. Byun, R. Myoung-Cheol, L. Seong-Whan, "Performance evaluation of face recognition algorithms on the asian face database, kfdb," in *Audio- and Video-Based Biometric Person Authentication*, J. Kittler, M. S. Nixon, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 557–565, ISBN: 978-3-540-44887-7.

[114] F. Solina, P. Peer, B. Batagelj, S. Juvan, J. Kovač, "Color-based face detection in the "15 seconds of fame" art installation," in *Proceedings of Mirage 2003, INRIA Rocquencourt*, France, Mar. 2003, pp. 37–47. [Online]. Available: http://www.lrv.fri.uni-lj.si/facedb.html.

[115] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, D. Zhao, "The cas-peal large-scale chinese face database and baseline evaluations," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 38, no. 1, pp. 149–161, Jan. 2008, ISSN: 1083-4427. DOI: 10.1109/TSMCA.2007.909557. [Online]. Available: http://www.jdl.ac.cn/peal/.

[116] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, "Overview of the face recognition grand challenge," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, IEEE, vol. 1, 2005, pp. 947–954.

[117] C. E. Thomaz, G. A. Giraldi, *A new ranking method for principal components analysis and its application to face image analysis*, 2010. [Online]. Available: https://fei.edu.br/~cet/facedatabase.html.

[118] L. Yin, X. Wei, Y. Sun, J. Wang, M. Rosato, "A 3d facial expression database for facial behavior research," in *International Conference on Automatic Face and Gesture Recognition (FGR06)*, Apr. 2006, pp. 211–216. DOI: 10.1109/FGR.2006.6. [Online]. Available: http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html.

[119] K. Ricanek, T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Apr. 2006, pp. 341–345. DOI: 10.1109/FGR.2006.78. [Online]. Available: https://ebill.uncw.edu/C20231_ustores/web/product_detail.jsp?PRODUCTID=8.

[120] G. B. Huang, M. Mattar, T. Berg, E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in *In Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*, Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, Marseille, France, Oct. 2008. [Online]. Available: http://vis-www.cs.umass.edu/lfw/.

[121] C. Zhong, Z. Sun, T. Tan, "Robust 3d face recognition using learned visual codebook," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp. 1–6. DOI: 10.1109/CVPR.2007.383279. [Online]. Available: http://www.cbsr.ia.ac.cn/english/3DFace%5C%20Databases.asp.

[122] S. Milborrow, J. Morkel, F. Nicolls, "The MUCT Landmarked Face Database," *Pattern Recognition Association of South Africa*, 2010. [Online]. Available: http://www.milbo.org/muct.

[123] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, L. Akarun, "Bosphorus database for 3d face analysis," in *Biometrics and Identity Management*, B. Schouten, N. C. Juul, A. Drygajlo, M. Tistarelli, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 47–56, ISBN: 978-3-540-89991-4. [Online]. Available: http://bosphorus.ee.boun.edu.tr/Home.aspx.

[124] R. Gross, I. Matthews, J. Cohn, T. Kanade, S. Baker, "Multi-pıe," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010, Best of Automatic Face and Gesture Recognition 2008, ISSN: 0262-8856. DOI: https://doi.org/10.1016/j.imavis.2009.08.002. [Online]. Available: http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html.

[125] X. Wang, X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2008.222. [Online]. Available: http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html.

[126] S. Z. Li, Z. L. and, "The hfb face database for heterogeneous face biometrics research," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2009, pp. 1–8. DOI: 10.1109/CVPRW.2009.5204149.

[127] X. Wang, X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2008.222. [Online]. Available: http://mmlab.ie.cuhk.edu.hk/archive/cufsf/.

[128] S. Z. Li, D. Yi, Z. Lei, S. Liao, "The casıa nır-vıs 2.0 face database," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2013, pp. 348–353. DOI: 10.1109/CVPRW.2013.59. [Online]. Available: https://pypi.org/project/bob.db.cbsr-nir-vis-2/.

[129] K. Chen, S. Gong, T. Xiang, C. C. Loy, "Cumulative attribute space for age and crowd density estimation," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2013, pp. 2467–2474. DOI: 10.1109/CVPR.2013.319. [Online]. Available: http://www-prima.inrialpes.fr/FGnet/html/benchmarks.html.

[130] W. A. Bainbridge, P. Isola, A. Oliva, "The intrinsic memorability of face photographs.," *Journal of experimental psychology. General*, vol. 142 4, pp. 1323–1334, 2013. [Online]. Available: https://www.wilmabainbridge.com/facememorability2.html.

[131] Y. Dong, L. Zhen, L. Shengcai, Z. L. Stan, "Learning face representation from scratch," *CoRR*, vol. abs/1411.7923, 2014. [Online]. Available: http://www.cbsr.ia.ac.cn/english/CASIA-WebFace-Database.html.

[132] B. Chen, C. Chen, W. H. Hsu, "Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset," *IEEE Transactions on Multimedia*, vol. 17, no. 6, pp. 804–815, Jun. 2015, ISSN: 1520-9210. DOI: 10.1109/TMM.2015.2420374. [Online]. Available: https://bcsiriuschen.github.io/CARC/.

[133] Y. Sun, Y. Chen, X. Wang, X. Tang, "Deep learning face representation by joint identification-verification," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2014, pp. 1988–1996. [Online]. Available: http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html.

[134] Y. Taigman, M. Yanga, M. Ranzato, L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708.

[135] F. Schroff, D. Kalenichenko, J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 815–823. DOI: 10.1109/CVPR.2015.7298682.

[136] Y. Guo, L. Zhang, Y. Hu, X. He, J. Gao, "Ms-celeb-1m: A dataset and benchmark for large-scale face recognition," *CoRR*, vol. abs/1607.08221, 2016. [Online]. Available: https://www.microsoft.com/en-us/research/publication/ms-celeb-1m-dataset-benchmark-large-scale-face-recognition-2/.

[137] A. Nech, I. Kemelmacher-Shlizerman, "Level playing field for million scale face recognition," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 3406–3415. DOI: 10.1109/CVPR.2017.363. [Online]. Available: http://megaface.cs.washington.edu/.

[138] S. Sengupta, J. Chen, C. Castillo, V. M. Patel, R. Chellappa, D. W. Jacobs, "Frontal to profile face verification in the wild," in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2016, pp. 1–9. DOI: 10.1109/WACV.2016.7477558. [Online]. Available: http://www.cfpw.io/.

[139] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age," *CoRR*, vol. abs/1710.08092, 2017. arXiv: 1710.08092. [Online]. Available: http://www.robots.ox.ac.uk/~vgg/data/vgg_face2/.

[140] T. Zheng, W. Deng, "Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments," Beijing University of Posts and Telecommunications, Tech. Rep. 18-01, Feb. 2018. [Online]. Available: http://www.whdeng.cn/CPLFW/index.html?reload=true.

[141] M. J. Black, Y. Yacoob, "Recognizing facial expressions in image sequences using local parameterized models of image motion," *International Journal of Computer Vision*, vol. 25, no. 1, pp. 23–48, 1997.

[142] K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre, "Xm2vtsdb: The extended m2vts database," in *Second International Conference on Audio and Video-based Biometric Person Authentication*, 1999, pp. 965–966. [Online]. Available: http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/.

[143] T. Kanade, J. F. Cohn, Y. Tian, "Comprehensive database for facial expression analysis," in *IEEE International Conference on Automatic Face and Gesture Recognition*, Mar. 2000, pp. 46–53. DOI: 10.1109/AFGR.2000.840611.

[144] R. Gross, J. Shi, "The cmu motion of body (mobo) database," Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-01-18, Jun. 2001. [Online]. Available: https://www.ri.cmu.edu/publications/the-cmu-motion-of-body-mobo-database/.

[145] K. Lee, J. Ho, M. Yang, D. Kriegman, "Visual tracking and recognition using probabilistic appearance manifolds," *Computer Vision and Image Understanding*, 2005. [Online]. Available: http://vision.ucsd.edu/~leekc/HondaUCSDVideoDatabase/HondaUCSD.html.

[146] A. J. O'Toole, J. Harms, S. L. Snow, D. R. Hurst, M. R. Pappas, J. H. Ayyad, H. Abdi, "A video database of moving faces and people," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 812–816, May 2005, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2005.90. [Online]. Available: http://www.utdallas.edu/dept/bbs/FACULTY_PAGES/otoole/database.htm.

[147] M. Breidt, D. W. Cunningham, C. Wallraven, *Max planck video database*. [Online]. Available: https://vdb.kyb.tuebingen.mpg.de/.

[148] E. Bailly-Bailliére, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, J.-P. Thiran, "The banca database and evaluation protocol," in *Proceedings of the 4th International Conference on Audio- and Video-based Biometric Person Authentication*, ser. AVBPA'03, Guildford, UK: Springer-Verlag, 2003, pp. 625–638, ISBN: 3-540-40302-7. [Online]. Available: http://www.ee.surrey.ac.uk/Research/VSSP/banca/.

[149] M. Kim, S. Kumar, V. Pavlovic, H. Rowley, "Face tracking and recognition with visual constraints in real-world videos," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2008, pp. 1–8. DOI: 10.1109/CVPR.2008.4587572.

[150] L. Yin, X. Chen, Y. Sun, T. Worm, M. Reale, "A high-resolution 3d dynamic facial expression database," in *2008 8th IEEE International Conference on Automatic Face Gesture Recognition*, Sep. 2008, pp. 1–6. DOI: 10.1109/AFGR.2008.4813324. [Online]. Available: http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html.

[151] C. McCool, S. Marcel, "Mobio database for the icpr 2010 face and speech competition," Idiap, Idiap-Com Idiap-Com-02-2009, Nov. 2009. [Online]. Available: https://www.idiap.ch/dataset/mobio.

[152] L. Wolf, T. Hassner, I. Maoz, "Face recognition in unconstrained videos with matched background similarity," in *IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, CO, USA, Jun. 2011, pp. 529–534. DOI: 10.1109/CVPR.2011.5995566. [Online]. Available: https://www.cs.tau.ac.il/~wolf/ytfaces/.

[153] ". R. Beveridge, P. J. Phillips, D. S. Bolme, B. A. Draper, G. H. Givens, Y. M. Lui, M. N. Teli, H. Zhang, W. T. Scruggs, K. W. Bowyer, P. J. Flynn, S. Cheng", "The challenge of face recognition from digital point-and-shoot cameras," in *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Arlington, VA, USA, Sep. 2013, pp. 1–8. DOI: 10.1109/BTAS.

2013.6712704. [Online]. Available: https://www.cs.colostate.edu/~vision/pasc/index.php.

[154] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, M. Burge, A. K. Jain, "Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1931–1939. DOI: 10.1109/CVPR.2015.7298803. [Online]. Available: https://www.nist.gov/itl/iad/image-group/ijb-dataset-request-form.

[155] C. Whitelam, E. Taborsky, A. Blanton, B. Maze, J. C. Adams, T. Miller, N. D. Kalka, A. K. Jain, J. A. Duncan, K. E. Allen, J. Cheney, P. Grother, "Iarpa janus benchmark-b face dataset," *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 592–600, 2017. [Online]. Available: https://www.nist.gov/programs-projects/face-challenges.

[156] A. Bansal, C. D. Castillo, R. Ranjan, R. Chellappa, "The do's and don'ts for cnn-based face verification," *CoRR*, vol. abs/1705.07426, 2017. arXiv: 1705.07426. [Online]. Available: http://umdfaces.io/.

[157] B. Maze, J. Adams, J. A. Duncan, N. Kalka, T. Miller, C. Otto, A. K. Jain, W. T. Niggel, J. Anderson, J. Cheney, P. Grother, "Iarpa janus benchmark - c: Face dataset and protocol," in *2018 International Conference on Biometrics (ICB)*, Feb. 2018, pp. 158–165. DOI: 10.1109/ICB2018.2018.00033. [Online]. Available: https://www.nist.gov/itl/iad/ig/ijb-c-dataset-request-form.

[158] S. Eberz, K. B. Rasmussen, V. Lenders, I. Martinovic, "Evaluating behavioral biometrics for continuous authentication: Challenges and metrics," in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 2017, pp. 386–399.

[159] N. Poh, C. Chan, J. Kittler, J. Fierrez, J. Galbally, "Description of metrics for the evaluation of biometric performance," *Biometrics Evaluation and Testing*, 2012.

[160] K. Messer, J. Kittler, M. Sadeghi, S. Marcel, C. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, J. Czyz, L. Vandendorpe, *et al.*, "Face verification competition on the xm2vts database," in *International Conference on Audio-and Video-Based Biometric Person Authentication*, Springer, 2003, pp. 964–974.

[161] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, "Overview of the face recognition grand challenge," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, IEEE, vol. 1, 2005, pp. 947–954.

[162] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, M. Bone, "Face recognition vendor test 2002," in *2003 IEEE International SOI Conference. Proceedings (Cat. No. 03CH37443)*, IEEE, 2003, p. 44.

[163] S. A. Rizvi, P. J. Phillips, H. Moon, "The feret verification testing protocol for face recognition algorithms," in *Proceedings third IEEE international conference on automatic face and gesture recognition*, IEEE, 1998, pp. 48–53.

[164] N. Erdogmus, S. Marcel, "Spoofing face recognition with 3d masks," *IEEE transactions on information forensics and security*, vol. 9, no. 7, pp. 1084–1097, 2014.

[165] R. Ramachandra, C. Busch, "Presentation attack detection methods for face recognition systems: A comprehensive survey," *ACM Computing Surveys (CSUR)*, vol. 50, no. 1, pp. 1–37, 2017.

[166] R. Shao, X. Lan, P. C. Yuen, "Joint discriminative learning of deep dynamic textures for 3d mask face anti-spoofing," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 4, pp. 923–938, 2018.

[167] M. Bagga, B. Singh, "Spoofing detection in face recognition: A review," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, IEEE, 2016, pp. 2037–2042.

[168] A. Ali, S. Hoque, F. Deravi, "Gaze stability for liveness detection," *Pattern Analysis and Applications*, vol. 21, no. 2, pp. 437–449, 2018.

[169] C.-H. Chu, Y.-K. Feng, "Study of eye blinking to improve face recognition for screen unlock on mobile devices," *Journal of Electrical Engineering and Technology*, vol. 13, no. 2, pp. 953–960, 2018.

[170] G. Lavrentyeva, O. Kudashev, A. Melnikov, M. De Marsico, Y. Matveev, "Interactive photo liveness for presentation attacks detection," in *International Conference Image Analysis and Recognition*, Springer, 2018, pp. 252–258.

[171] G. Lenz, S.-H. Ieng, R. Benosman, "High speed event-based face detection and tracking in the blink of an eye," *arXiv preprint arXiv:1803.10106*, 2018.

[172] M. Singh, A. Arora, "A novel face liveness detection algorithm with multiple liveness indicators," *Wireless Personal Communications*, vol. 100, no. 4, pp. 1677–1687, 2018.

[173] D. K. Wong, R. Janakiraman, *Face liveness detection*, US Patent App. 15/610,273, Dec. 2018.

[174] G. Chetty, "Robust audio visual biometric person authentication with liveness verification," in *Intelligent Multimedia Analysis for Security Applications*, Springer, 2010, pp. 59–78.

[175] S. M. Hatture, P. Karchi, "Prevention of spoof attack in biometric system using liveness detection," *Int. J. Latest Trends Eng. Technol*, pp. 42–49, 2013.

[176] C. Kant, N. Sharma, "Fake face recognition using fusion of thermal imaging and skin elasticity," *IJCSC*, vol. 4, pp. 65–72, 2013.

[177] K. Kollreider, H. Fronthaler, M. I. Faraj, J. Bigun, "Real-time face detection and motion analysis with application in "liveness" assessment," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 548–558, 2007.

[178] A. Lagorio, M. Tistarelli, M. Cadoni, C. Fookes, S. Sridharan, "Liveness detection based on 3d face shape analysis," in *2013 International Workshop on Biometrics and Forensics (IWBF)*, IEEE, 2013, pp. 1–4.

[179] E. M. Nowara, A. Sabharwal, A. Veeraraghavan, "Ppgsecure: Biometric presentation attack detection using photopletysmograms," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, IEEE, 2017, pp. 56–62.

[180] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, Z. J. Wang, "Video-based heart rate measurement: Recent advances and future prospects," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3600–3615, 2018.

[181] H. Demirezen, C. E. Erdem, "Remote photoplethysmography using nonlinear mode decomposition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2018, pp. 1060–1064.

[182] M. Killioğlu, M. Taşkiran, N. Kahraman, "Anti-spoofing in face recognition with liveness detection using pupil tracking," in *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, IEEE, 2017, pp. 000 087–000 092.

[183] K. W. Bowyer, K. Chang, P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition," *Computer vision and image understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[184] B. M. Chrzan, "Liveness detection for face recognition," *Master's thesis*, 2014.

[185] J. Galbally, S. Marcel, J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[186] A. Hadid, "Face biometrics under spoofing attacks: Vulnerabilities, countermeasures, open issues, and research directions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 113–118.

[187] K. A. Nixon, V. Aimale, R. K. Rowe, "Spoof detection schemes," in *Handbook of biometrics*, Springer, 2008, pp. 403–423.

[188] G. Pan, Z. Wu, L. Sun, "Liveness detection for face recognition," *Recent advances in face recognition*, pp. 109–124, 2008.

[189] H. Li, P. He, S. Wang, A. Rocha, X. Jiang, A. C. Kot, "Learning generalized deep feature representation for face anti-spoofing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 10, pp. 2639–2652, 2018.

[190] C. Nagpal, S. R. Dubey, "A performance evaluation of convolutional neural networks for face anti spoofing," in *2019 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2019, pp. 1–8.

[191] Y. Liu, J. Stehouwer, A. Jourabloo, X. Liu, "Deep tree learning for zero-shot face anti-spoofing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4680–4689.

[192] X. Yang, W. Luo, L. Bao, Y. Gao, D. Gong, S. Zheng, Z. Li, W. Liu, "Face anti-spoofing: Model matters, so does data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3507–3516.

[193] S. Zhang, X. Wang, A. Liu, C. Zhao, J. Wan, S. Escalera, H. Shi, Z. Wang, S. Z. Li, "A dataset and benchmark for large-scale multi-modal face anti-spoofing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 919–928.

[194] A. Liu, J. Wan, S. Escalera, H. Jair Escalante, Z. Tan, Q. Yuan, K. Wang, C. Lin, G. Guo, I. Guyon, *et al.*, "Multi-modal face anti-spoofing attack detection challenge at cvpr2019," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.

[195] Y. Sun, *Deep learning face representation by joint identification-verification*. The Chinese University of Hong Kong (Hong Kong), 2015.

[196] Y. Sun, X. Wang, X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1891–1898.

[197] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "Web-scale training for face identification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2746–2754.

[198] M. Mehdipour Ghazi, H. Kemal Ekenel, "A comprehensive analysis of deep learning based representation for face recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2016, pp. 34–41.

[199] S. Xie, Z. Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1395–1403.

[200] J. Yang, S. E. Reed, M.-H. Yang, H. Lee, "Weakly-supervised disentangling with recurrent transformations for 3d view synthesis," *Advances in neural information processing systems*, vol. 28, pp. 1099–1107, 2015.

[201] I. Masi, T. Hassner, A. T. Tran, G. Medioni, "Rapid synthesis of massive face sets for improved face recognition," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, IEEE, 2017, pp. 604–611.

[202] I. Masi, A. T. Trn, T. Hassner, J. T. Leksut, G. Medioni, "Do we really need to collect millions of faces for effective face recognition?" İn *European conference on computer vision*, Springer, 2016, pp. 579–596.

[203] Y. Qian, W. Deng, J. Hu, "Task specific networks for identity and face variation," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, 2018, pp. 271–277.

[204] J. Yim, H. Jung, B. Yoo, C. Choi, D. Park, J. Kim, "Rotating your face using multi-task deep neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 676–684.

[205] J. Bao, D. Chen, F. Wen, H. Li, G. Hua, "Towards open-set identity preserving face synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6713–6722.

[206] J. Bao, D. Chen, F. Wen, H. Li, G. Hua, "Cvae-gan: Fine-grained image generation through asymmetric training," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2745–2754.

[207] W. Chai, W. Deng, H. Shen, "Cross-generating gan for facial identity preserving," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, 2018, pp. 130–134.

[208] Y. Shen, P. Luo, J. Yan, X. Wang, X. Tang, "Faceid-gan: Learning a symmetry three-player gan for identity-preserving face synthesis," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 821–830.

[209] J. Zhao, L. Xiong, J. Karlekar, J. Li, F. Zhao, Z. Wang, S. Pranata, S. Shen, S. Yan, J. Feng, "Dual-agent gans for photorealistic and ıdentity preserving profile face synthesis.," in *NIPS*, vol. 2, 2017, p. 3.

[210] Z. Zhu, P. Luo, X. Wang, X. Tang, "Deep learning identity-preserving face space," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 113–120.

[211] L. Tran, X. Yin, X. Liu, "Disentangled representation learning gan for pose-invariant face recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1415–1424.

[212] Z. Zhang, X. Chen, B. Wang, G. Hu, W. Zuo, E. R. Hancock, "Face frontalization using an appearance-flow-based convolutional neural network," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2187–2199, 2018.

[213] Z. An, W. Deng, J. Hu, Y. Zhong, Y. Zhao, "Apa: Adaptive pose alignment for pose-invariant face recognition," *IEEE Access*, vol. 7, pp. 14 653–14 670, 2019.

[214] W. Zhang, X. Zhao, J.-M. Morvan, L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 3, pp. 611–624, 2018.

[215] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[216] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[217] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[218] J. Hu, L. Shen, G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.

[219] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[220] E. Hand, R. Chellappa, "Attributes for improved attributes: A multi-task network utilizing implicit and explicit relationships for facial attribute classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.

[221] E. P. Xing, A. Y. Ng, M. I. Jordan, S. Russell, "Distance metric learning with application to clustering with side-information," in *NIPS*, Citeseer, vol. 15, 2002, p. 12.

[222] K. Q. Weinberger, L. K. Saul, "Distance metric learning for large margin nearest neighbor classification.," *Journal of machine learning research*, vol. 10, no. 2, 2009.

[223] W. Liu, Y. Wen, Z. Yu, M. Yang, "Large-margin softmax loss for convolutional neural networks.," in *ICML*, vol. 2, 2016, p. 7.

[224] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, L. Song, "Sphereface: Deep hypersphere embedding for face recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 212–220.

[225] W. Liu, Y.-M. Zhang, X. Li, Z. Yu, B. Dai, T. Zhao, L. Song, "Deep hyperspherical learning," *arXiv preprint arXiv:1711.03189*, 2017.

[226] Z. Ding, Y. Guo, L. Zhang, Y. Fu, "One-shot face recognition via generative learning," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, 2018, pp. 1–7.

[227] Y. Guo, L. Zhang, "One-shot face recognition by promoting underrepresented classes," *arXiv preprint arXiv:1707.05574*, 2017.

[228] A. Jadhav, V. P. Namboodiri, K. Venkatesh, "Deep attributes for one-shot face recognition," in *European Conference on Computer Vision*, Springer, 2016, pp. 516–523.

[229] L. Wang, Y. Li, S. Wang, "Feature learning for one-shot face recognition," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2018, pp. 2386–2390.

[230] Y. Wu, H. Liu, Y. Fu, "Low-shot face recognition with hybrid classifiers," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 1933–1939.

[231] Y. Sun, D. Liang, X. Wang, X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint arXiv:1502.00873*, 2015.

[232] Y. Wen, K. Zhang, Z. Li, Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *European conference on computer vision*, Springer, 2016, pp. 499–515.

[233] J. Deng, Y. Zhou, S. Zafeiriou, "Marginal loss for deep face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 60–68.

[234] M. Hasnat, J. Bohné, J. Milgram, S. Gentric, L. Chen, *et al.*, "Von mises-fisher mixture model-based deep learning: Application to face verification," *arXiv preprint arXiv:1706.04264*, 2017.

[235] Y. Liu, H. Li, X. Wang, "Rethinking feature discrimination and polymerization for large-scale recognition," *arXiv preprint arXiv:1710.00870*, 2017.

[236] R. Ranjan, C. D. Castillo, R. Chellappa, "L2-constrained softmax loss for discriminative face verification," *arXiv preprint arXiv:1703.09507*, 2017.

[237] F. Wang, X. Xiang, J. Cheng, A. L. Yuille, "Normface: L2 hypersphere embedding for face verification," in *Proceedings of the 25th ACM international conference on Multimedia*, 2017, pp. 1041–1049.

[238] X. Zhang, Z. Fang, Y. Wen, Z. Li, Y. Qiao, "Range loss for deep face recognition with long-tailed training data," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5409–5418.

[239] A. Hadid, M. Pietikäinen, "An experimental investigation about the integration of facial dynamics in video-based face recognition," *ELCVIA: electronic letters on computer vision and image analysis*, pp. 1–13, 2005.

[240] J. F. Cohn, K. Schmidt, R. Gross, P. Ekman, "Individual differences in facial expression: Stability over time, relation to self-reported emotion, and ability to inform person identification," in *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, IEEE, 2002, pp. 491–496.

[241] X. Liu, T. Cheng, "Video-based face recognition using adaptive hidden markov models," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, IEEE, vol. 1, 2003, pp. I–I.

[242] F. Matta, J.-l. Dugelay, "A behavioural approach to person recognition," in *2006 IEEE International Conference on Multimedia and Expo*, IEEE, 2006, pp. 1461–1464.

[243] F. Matta, J.-L. Dugelay, "Video face recognition: A physiological and behavioural multimodal approach," in *2007 IEEE International Conference on Image Processing*, IEEE, vol. 4, 2007, pp. IV–497.

[244] S. Zafeiriou, M. Pantic, "Facial behaviometrics: The case of facial deformation in spontaneous smile/laughter," in *CVPR 2011 WORKSHOPS*, IEEE, 2011, pp. 13–19.

[245] K. L. Schmidt, J. F. Cohn, "Dynamics of facial expression: Normative characteristics and individual differences," in *IEEE International Conference on Multimedia and Expo, 2001. ICME 2001.*, Citeseer, 2001, pp. 547–550.

[246] A. L. Kashyap, S. Tulyakov, V. Govindaraju, "Facial behavior as a soft biometric," in *2012 5th IAPR International Conference on Biometrics (ICB)*, IEEE, 2012, pp. 147–151.

[247] M. Shreve, E. A. Bernal, Q. Li, J. Kumar, R. Bala, "A study on the discriminability of facs from spontaneous facial expressions," in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, pp. 1674–1678.

[248] L.-F. Chen, H.-Y. Liao, J.-C. Lin, "Person identification using facial motion," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, IEEE, vol. 2, 2001, pp. 677–680.

[249] M.-I. Faraj, J. Bigun, "Audio–visual person authentication using lip-motion from orientation maps," *Pattern recognition letters*, vol. 28, no. 11, pp. 1368–1382, 2007.

[250] P. Tsai, T. Hintz, T. Jan, "Facial behavior as behavior biometric? an empirical study," in *2007 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2007, pp. 3917–3922.

[251] R. E. Haamer, K. Kulkarni, N. Imanpour, M. A. Haque, E. Avots, M. Breisch, K. Nasrollahi, S. Escalera, C. Ozcinar, X. Baro, *et al.*, "Changes in facial expression as biometric: A database and benchmarks of identification," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, 2018, pp. 621–628.

[252] M. Bicego, E. Grosso, M. Tistarelli, "Person authentication from video of faces: A behavioral and physiological approach using pseudo hierarchical hidden markov models," in *International Conference on Biometrics*, Springer, 2006, pp. 113–120.

[253] M. Tistarelli, M. Bicego, E. Grosso, "Dynamic face recognition: From human to machine vision," *Image and Vision Computing*, vol. 27, no. 3, pp. 222–232, 2009.

[254] A. Asthana, S. Zafeiriou, S. Cheng, M. Pantic, "Incremental face alignment in the wild," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1859–1866.

[255] X. Zhu, D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *2012 IEEE conference on computer vision and pattern recognition*, IEEE, 2012, pp. 2879–2886.

[256] P. Ekman, W. V. Friesen, "Felt, false, and miserable smiles," *Journal of nonverbal behavior*, vol. 6, no. 4, pp. 238–252, 1982.

[257] P. Geurts, D. Ernst, L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3–42, 2006.

[258] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Ieee, vol. 1, 2005, pp. 886–893.

[259] D. E. King, "Max-margin object detection," *arXiv preprint arXiv:1502.00046*, 2015.

[260] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "Closing the gap to human-level performance in face verification. deepface," in *IEEE Computer Vision and Pattern Recognition (CVPR)*, vol. 5, 2014, p. 6.

[261] J. Lu, G. Wang, J. Zhou, "Simultaneous feature and dictionary learning for image set based face recognition," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 4042–4054, 2017.

[262] Y. Rao, J. Lu, J. Zhou, "Attention-aware deep reinforcement learning for video face recognition," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 3931–3940.

[263] J. Deng, J. Guo, N. Xue, S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4690–4699.

[264] M. Loukadakis, J. Cano, M. O'Boyle, "Accelerating deep neural networks on low power heterogeneous architectures," 2018.

[265] C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)*, 2007.

[266] E. Fix, J. L. Hodges Jr, "Discriminatory analysis-nonparametric discrimination: Small sample performance," CALIFORNIA UNIV BERKELEY, Tech. Rep., 1952.

[267] F. Wallhoff, B. Schuller, M. Hawellek, G. Rigoll, "Efficient recognition of authentic dynamic facial expressions on the feedtum database," in *2006 IEEE international conference on multimedia and expo*, IEEE, 2006, pp. 493–496.

[268] R. Yan, L. Shao, "Blind image blur estimation via deep learning," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1910–1921, 2016.

[269] J. H. Chuah, H. Y. Khaw, F. C. Soon, C.-O. Chow, "Detection of gaussian noise and its level using deep convolutional neural network," in *TENCON 2017-2017 IEEE Region 10 Conference*, IEEE, 2017, pp. 2447–2450.

[270] J. Das, B. Das, J. Saikia, S. Nirmala, “Removal of salt and pepper noise using selective adaptive median filter,” in *2016 International Conference on Accessibility to Digital World (ICADW)*, IEEE, 2016, pp. 203–206.

TEZDEN ÜRETİLMİŞ YAYINLAR

Makale

1. Taskiran, Murat, Nihan Kahraman, and Cigdem Eroglu Erdem. "Face recognition: Past, present and future (a review)." Digital Signal Processing (2020): 102809. (SCI Expanded İndekslerine Giren Dergi)

2. Taskiran, Murat, Nihan Kahraman, and Cigdem Eroglu Erdem. "Hybrid face recognition under adverse conditions using appearance-based and dynamic features of smile expression." IET Biometrics 10.1 (2021): 99-115. (SCI Expanded İndekslerine Giren Dergi)

Konferans Bildirisi

1. Taskiran, Murat, et al. "Face recognition using dynamic features extracted from smile videos." 2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA). IEEE, 2019.

Proje

1. "Çok Kipli Doğal Yüz İfadesi Videolarından Yumuşak Biyometrik Veri Çıkarılması", TÜBİTAK Destekli Projeler (TÜBİTAK 1001) No: EEAG-116E088.