T.C.
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

KIZILÖTESİ GÖRÜNTÜLERDE DERİN ÖĞRENME TABANLI YÜZ TESPİTİ İÇİN YENİ SINIRLAYICI KUTU BELİRLEME YÖNTEMLERİ

Hazırlayan
Ahmet ÜNLÜHİSARCIKLI

Danışman
Dr. Öğr. Üyesi Ahmet Nusret TOPRAK

Yüksek Lisans Tezi

Ocak 2024
KAYSERİ

T.C.
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

KIZILÖTESİ GÖRÜNTÜLERDE DERİN ÖĞRENME TABANLI YÜZ TESPİTİ İÇİN YENİ SINIRLAYICI KUTU BELİRLEME YÖNTEMLERİ

(Yüksek Lisans Tezi)

Hazırlayan
Ahmet ÜNLÜHİSARCIKLI

Danışman
Dr. Öğr. Üyesi Ahmet Nusret TOPRAK

Bu çalışma, Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından FYL-2023-12497 kodlu proje ile desteklenmiştir.

Ocak 2024
KAYSERİ

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Ahmet ÜNLÜHİSARCIKLI

“Kızılötesi Görüntülerde Derin Öğrenme Tabanlı Yüz Tespiti İçin Yeni Sınırlayıcı Kutu Belirleme Yöntemleri” adlı Yüksek Lisans tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ne uygun olarak hazırlanmıştır.

Tezi Hazırlayan	**Danışman**
Ahmet ÜNLÜHİSARCIKLI	Dr. Öğr. Üyesi Ahmet Nusret TOPRAK

Bilgisayar Mühendisliği ABD Başkanı
Prof. Dr. Veysel ASLANTAŞ

Dr. Öğr. Üyesi Ahmet Nusret TOPRAK danışmanlığında **Ahmet ÜNLÜHİSARCIKLI** tarafından hazırlanan **"Kızılötesi Görüntülerde Derin Öğrenme Tabanlı Yüz Tespiti İçin Yeni Sınırlayıcı Kutu Belirleme Yöntemleri"** adlı bu çalışma, jürimiz tarafından Erciyes Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalında **Yüksek Lisans** tezi olarak kabul edilmiştir.

JÜRİ :

Danışman : Dr. Öğr. Üyesi Ahmet Nusret TOPRAK

Üye : Dr. Öğr. Üyesi Ömer Faruk SARITAŞ

Üye : Dr. Öğr. Üyesi Tayyip ÖZCAN

ONAY :

Bu tezin kabülü Enstitü Yönetim Kurulunun ..... tarih ve ..... sayılı kararı ile onaylanmıştır.

Prof. Dr. Emel KIZILKAYA AYDOĞAN
Enstitü Müdürü

TEŞEKKÜR

Tez çalışmam boyunca beni yönlendiren ve değerli yardımlarını esirgemeyen tez danışmanım Dr. Öğr. Üyesi Ahmet Nusret TOPRAK'a teşekkür ederim.

Bu tez çalışmasına, FYL-2023-12497 numaralı proje kapsamında destek sağlayan Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi'ne teşekkür ederim.

Eğitim-öğretim hayatım boyunca desteklerini esirgemeyen, her zaman yanımda olup her türlü desteği sağlayan anneme, babama, kardeşime, anneanneme ve Rümeysa'ya sonsuz teşekkürlerimi sunarım.

KIZILÖTESİ GÖRÜNTÜLERDE DERİN ÖĞRENME TABANLI YÜZ TESPİTİ İÇİN YENİ SINIRLAYICI KUTU BELİRLEME YÖNTEMLERİ

Ahmet ÜNLÜHİSARCIKLI
Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Yüksek Lisans Tezi, Ocak 2024
Danışman : Dr. Öğr. Üyesi Ahmet Nusret TOPRAK

ÖZET

Yüz tespiti; güvenlik, sağlık, endüstri ve biyometri gibi alanlarda kritik rol oynamaktadır. Bahsi geçen alanlarda, aydınlatma durumundan bağımsız olarak yüz tespitinin doğru ve verimli bir şekilde gerçekleştirilebilmesi büyük öneme sahiptir. Ancak, yetersiz aydınlatma koşullarında verimli bir şekilde yüz tespiti yapmak zor olabilmektedir. Bu problem doğrultusunda, yetersiz aydınlatma koşullarında dahi verimli çalışabilen, yakın kızılötesi görüntüler üzerinde yüz tespitine odaklanan bir tez çalışması sunulmaktadır. Bu amaçla, öncelikle literatürde iyi bilinen dokuz farklı yüz tespiti derin sinir ağı modeli belirlenmiş ve yakın kızılötesi görüntülerdeki başarımlarını ortaya koymak amacıyla görsel ve sayısal olarak karşılaştırılmıştır. Buna ilaveten, çalışma kapsamında önerilen yaklaşım birden fazla yüz tespiti modelinin ürettiği sınırlayıcı kutu topluluklarının beraber analiz edilmesi ile daha başarılı ve daha kapsayıcı yeni sınırlayıcı kutular üretilebileceği ilkesine dayanmaktadır. Buradan hareketle, Kombinasyonel Oylama ve Maksimum Olmayan Ortalama adları verilen iki yeni sınırlayıcı kutu belirleme yöntemi önerilmektedir. Önerilen yöntemler literatürdeki sınırlayıcı kutu belirleme yöntemleri ile karşılaştırılmış ve önerilen yöntemlerden Kombinasyonel Oylama'nın, literatürdeki sınırlayıcı kutu belirleme yöntemlerinden daha başarılı sonuçlar ortaya koyduğu görülmüştür.

Anahtar Kelimeler: Yüz tespiti, yakın kızılötesi görüntüler, sınırlayıcı kutu topluluğu birleştirme

NEW BOUNDING BOX ENSEMBLE METHODS FOR DEEP LEARNING BASED FACE DETECTION IN INFRARED IMAGES

Ahmet ÜNLÜHİSARCIKLI
Erciyes University, Graduate School of Natural and Applied Sciences
M.Sc. Thesis, January 2024
Supervisor: Asst. Prof. Ahmet Nusret TOPRAK

ABSTRACT

Face detection plays a crucial role in various areas such as security, healthcare, industry, bio-metrics, etc. It is essential to perform face detection accurately and efficiently in the mentioned areas regardless of the lighting conditions. However, detecting faces in poor lighting conditions can be challenging. To address this issue, this study presents an approach focusing on face detection on near infrared images, which can work efficiently even in insufficient lighting conditions. For this purpose, first of all, nine different state-of-the-art face detection models in the literature were determined and compared visually and numerically to reveal their performance in near infrared images. In addition, the approach proposed in this study is based on the principle that a more successful and inclusive bounding-box can be produced by using a bounding-box ensemble generated by more than one face detection model. Based on this, two new bounding-box ensemble methods called Combinational Voting and Non-Maximum Averaging are proposed. The proposed methods are compared with bounding-box ensemble methods in the literature and Combinational Voting produces more successful results than the other methods in the literature.

Keywords: Face detection, near-infrared images, bounding-box ensemble

İÇİNDEKİLER

KIZILÖTESİ GÖRÜNTÜLERDE DERİN ÖĞRENME TABANLI YÜZ TESPİTİ İÇİN YENİ SINIRLAYICI KUTU BELİRLEME YÖNTEMLERİ

BİLİMSEL ETİĞE UYGUNLUK SAYFASI ii

YÖNERGEYE UYGUNLUK SAYFASI iii

KABUL VE ONAY iv

TEŞEKKÜR v

ÖZET vi

ABSTRACT vii

İÇİNDEKİLER viii

TABLOLAR LİSTESİ x

ŞEKİLLER LİSTESİ xi

GİRİŞ 1

1. BÖLÜM
GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

1.1. Yüz Tespiti 5

1.1.1. Yetersiz Aydınlatma Şartlarında Yüz Tespiti 5

1.1.2. Yakın Kızılötesi Görüntülerde Yüz Tespiti 7

1.2. Yüz Tespit Yöntemleri 10

1.2.1. Dlib 10

1.2.2. FaceBoxes 11

1.2.3. DSFD 12

1.2.4. BlazeFace 13

1.2.5. CenterFace 15

1.2.6. RetinaFace 16

1.2.7. SCRFD 17

1.2.8. YuNet 18

1.3. Sınırlayıcı Kutu Belirleme Yöntemleri 20

1.3.1. Maksimum Olmayan Bastırma 21

1.3.2. Maksimum Olmayan Ağırlıklandırma . . . 22

1.3.3. Ağırlıklı Kutular Birleşimi . . . 23

1.3.4. Ortalama Alma . . . 24

2. BÖLÜM
YÖNTEM VE MATERYAL

2.1. TUFTS Veri Seti . . . 27

2.2. Sınırlayıcı Kutu Belirleme Yöntemleri Kullanarak Yüz Tespiti . . . 28

2.2.1. Maksimum Olmayan Ortalama . . . 30

2.2.2. Kombinasyonel Oylama . . . 31

3. BÖLÜM
BULGULAR

3.1. Deney Altyapısı . . . 35

3.2. Performans Değerlendirme Metrikleri . . . 35

3.3. Yüz Tespiti Derin Sinir Ağı Modellerinin Karşılaştırılması . . . 38

3.3.1. Yüz Tespiti Modellerinin Görsel Karşılaştırması . . . 38

3.3.2. Yüz Tespiti Modellerinin Sayısal Karşılaştırması . . . 44

3.3.3. Yüz Tespiti Modellerinin İşlem Zamanlarının Karşılaştırması . . . 51

3.4. Sınırlayıcı Kutu Belirleme Yöntemlerinin Karşılaştırılması . . . 52

3.4.1. Yöntemlerin Görsel Karşılaştırması . . . 52

3.4.2. Yöntemlerin Sayısal Karşılaştırması . . . 57

4. BÖLÜM
TARTIŞMA, SONUÇ VE ÖNERİLER

4.1. Tartışma . . . 59

4.2. Sonuç ve Öneriler . . . 60

KAYNAKÇA . . . 62

EKLER . . . 67

ÖZGEÇMİŞ . . . 68

TABLOLAR LİSTESİ

Tablo 3.1. Yüz tespit yöntemlerinin düz kafa pozisyonunda karşılaştırılması . . 44

Tablo 3.2. Yüz tespit yöntemlerinin yan kafa pozisyonunda karşılaştırılması . . 45

Tablo 3.3. Yüz tespit yöntemlerinin bütün kafa pozisyonlarında karşılaştırılması 45

Tablo 3.4. İşlem zamanı karşılaştırma tablosu 51

Tablo 3.5. Modellere ait omurga ağı, parametre ve FLOP karşılaştırma tablosu . 52

Tablo 3.6. En iyi sonucu veren beş modelin (SCRFD, RetinaFace, DSFD, CenterFace, FaceBoxes) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması 57

Tablo 3.7. En iyi sonucu veren dört modelin (SCRFD, RetinaFace, DSFD, CenterFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması 58

Tablo 3.8. En iyi sonucu veren üç modelin (SCRFD, RetinaFace, DSFD) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması 58

Tablo 3.9. En iyi sonucu veren iki modelin (SCRFD, RetinaFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması 58

ŞEKİLLER LİSTESİ

Şekil 1.1. Işık tayfı 8

Şekil 1.2. FaceBoxes mimarisi 11

Şekil 1.3. DSFD mimarisi 13

Şekil 1.4. BlazeFace mimarisi 14

Şekil 1.5. CenterFace mimarisi 15

Şekil 1.6. RetinaFace mimarisi 17

Şekil 1.7. YuNet mimarisi 19

Şekil 1.8. Ortalama alma yöntemi çalışma şeması 24

Şekil 2.1. TUFTS veri seti yakın kızılötesi örnek görüntüler 27

Şekil 2.2. Sınırlayıcı kutu belirleme yöntemi akış diyagramı 28

Şekil 2.3. NMS ve NMA yöntemleri karşılaştırma şeması 30

Şekil 2.4. Kombinasyonel oylama yöntemi çalışma şeması 32

Şekil 3.1. Karmaşıklık matrisi 35

Şekil 3.2. Örnek yüz tespiti: (a) gerçek yüz bölgesi, (b) tahmin edilen yüz bölgesi, (c) yapılan tahmin için TP, FN, FP ve TN alanlar 36

Şekil 3.3. Farklı yüz tespit yöntemleri tarafından tespit edilen yüz bölgeleri . . . 39

Şekil 3.4. Yüz tespit yöntemleri tarafından bulunan yüz bölgeleri ve ürettikleri güven skorları 40

Şekil 3.5. Referans değer ve CenterFace tarafından çıkarılan yüz bölgeleri . . . 40

Şekil 3.6. Kullanılan bütün yüz tespit yöntemleri tarafından aynı görüntüde yüz tespiti: (a) referans yüz bölgesi, (b) SCRFD, (c) RetinaFace, (d) DSFD, (e) CenterFace, (f) FaceBoxes, (g) RetinaFace(MobileNet), (h) BlazeFace, (k) Dlib, (l) YuNet 43

Şekil 3.7. Düz kafa pozisyonunda İşlem Karakteristiği Eğrisi 46

Şekil 3.8. Yan kafa pozisyonunda İşlem Karakteristiği Eğrisi 47

Şekil 3.9. Bütün kafa pozisyonlarında İşlem Karakteristiği Eğrisi 48

Şekil 3.10. Düz kafa pozisyonunda Kesinlik Duyarlılık Eğrisi 49

Şekil 3.11. Yan kafa pozisyonunda Kesinlik Duyarlılık Eğrisi 50

Şekil 3.12. Bütün kafa pozisyonlarında Kesinlik Duyarlılık Eğrisi 50

Şekil 3.13. En iyi sonucu veren modellerin kullanımı ile Kombinasyonel Oylama yönteminden yeni sınırlayıcı kutu üretilmesi: (a) beş model, (b) dört model, (c) üç model, (d) iki model 54

Şekil 3.14. En iyi sonucu veren modellerin kullanımı ile Maksimum Olmayan Ortalama yönteminden yeni sınırlayıcı kutu üretilmesi: (a) beş model, (b) dört model, (c) üç model, (d) iki model 56

GİRİŞ

Yüz tespiti yöntemleri, sayısal görüntüler ile videolarda bulunan yüzleri tanımlamak ve konumlarını belirlemek için genellikle makine öğrenimi ve derin öğrenme tekniklerinden yararlanan yöntemlerdir. Bu yöntemler, verilen görüntüde yüz olup olmadığını ve varsa bu yüzlerin konumlarını tespit etmeyi amaçlamaktadır.

Yüz tespiti; güvenlik, sağlık, endüstri ve biyometik analizlerin kullanıldığı diğer bir çok sektörde kritik öneme sahiptir [1]. Bu sebeple, yüz tespiti işlemlerinin aydınlatma durumundan bağımsız bir şekilde gerçekleştirilmesi önem arz etmektedir. Ancak özellikle gece vakitleri gibi yetersiz aydınlatma şartlarında yüz tespiti işlemlerinin başarısı azalmaktadır. Diğer taraftan, tespit edilen bu probleme ilişkin literatürde yapılmış çok az sayıda çalışma bulunmaktadır. Bu değerlendirmeler doğrultusunda, yetersiz aydınlatma şartlarında dahi görüntü oluşumuna imkan sağlayan yakın kızılötesi görüntülerin kullanımı ile yüz tespiti yöntemlerinin başarımları analizi edilmiştir.

Tez çalışması kapsamında, güncel yüz tespiti yöntemlerinin yakın kızılötesi görüntüler üzerindeki başarımlarının araştırılması ve yeni yaklaşımlar önerilerek başarımlarının artırılması hedeflenmektedir.Ancak bu amaç doğrultusunda kullanılabilecek etiketli bir veri seti literatürde bulunmamaktadır. Bu sebeple; öncelikle TUFTS veri setinde [2] yer alan yakın kızılötesi görüntüler etiketlenmiştir. Buradan elde edilen veriler, yapılan çalışmalar kapsamında referans değer olarak kullanılmıştır.

Literatürde; yüz tespiti yapmak için önerilmiş yöntemler bulunmaktadır. Bu yöntemler arasında, derin öğrenme kullananların daha başarılı sonuçlar ürettikleri bilinmektedir [3]. Bu sebeple, literatürde başarımları bilinen güncel yüz tespiti derin sinir ağı modelleri belirlenmiştir. Belirlenen modellerin yakın kızılötesi görüntüler üzerindeki başarımları analiz edilmiştir. Karşılaştırmalı olarak yapılan bu analiz işlemleri neticesinde ilgili modellere ait somut performans değerleri elde edilmiştir.

Sınırlayıcı kutu belirleme yöntemleri, girdi olarak verilen sınırlayıcı kutuların beraber analiz edilmesi ile daha başarılı tespit alanları elde etmeyi amaçlamaktadır. Bu bağlamda, belirlenen yüz tespiti modellerinin sınırlayıcı kutu belirleme yöntemlerine girdi olarak verilmesi ile daha başarılı yüz bölgesi tespiti yapılması hedeflenmiştir. Hedef doğrultusunda, dört farklı sınırlayıcı kutu belirleme yöntemi kullanılmış ve iki yeni sınırlayıcı kutu belirleme yöntemi de sunulan tez çalışması kapsamında önerilmiştir.

Bu tez çalışmasının katkıları aşağıdaki şekilde özetlenebilir:

1. TUFTS veri seti içerisinde yer alan bütün yakın kızılötesi görüntülerdeki yüz konumları etiketlenmiştir.
2. Literatürde başarımları bilinen güncel yüz tespiti yöntemlerinin yakın kızılötesi görüntüler üzerindeki performansları karşılaştırmalı olarak analiz edilmiştir.
3. Kullanılan yüz tespit yöntemlerinin yakın kızılötesi görüntülerde bulmuş oldukları yüz konumlarını belirleyen sınırlayıcı kutuların iyileştirilmesi için literatürdeki güncel sınırlayıcı kutu belirleme yöntemlerinin başarımları karşılaştırılmıştır.
4. Kombinasyonel Oylama ve Maksimum Olmayan Ortalama adları verilen iki yeni sınırlayıcı kutu belirleme yöntemi önerilmiştir.

Tezin Motivasyonu

Yüz tespiti; başta güvenlik sektörü olmak üzere biyometrik verilerin kullanıldığı hemen hemen her sektörde kritik rol oynamaktadır. Bu nedenle, yüz tespitinin aydınlatma durumundan bağımsız olarak yapılabilmesi gerekmektedir. Buna dayanarak, yetersiz aydınlatma şartlarında dahi çalışabilen, yakın kızılötesi görüntüler üzerinde yüz tespitine odaklanan bir tez çalışması yürütülmüştür. Yürütülen çalışmanın konusu dahilinde, literatürde çok az sayıda çalışma bulunmaktadır. Bu sebeple, ilgili çalışma literatüre katkı sağlayacak niteliktedir.

Tezin Organizasyonu

Birinci bölümde; yüz tespitinin kullanım biçimi ve öneminden bahsedilmektedir. Yüz tespiti, yetersiz aydınlatma şartlarında ve yakın kızılötesi görüntülerde olmak üzere iki alt başlık altında ele alıarak literatür taramaları verilmektedir. Yakın

kızılötesi görüntüler üzerinde başarımlarının araştırılacağı yüz tespit yöntemleri ve bu yöntemlerin başarımlarının artırılması hedefiyle seçilen sınırlayıcı kutu belirleme yöntemleri tanıtılmaktadır.

İkinci bölümde; tez çalışması kapsamında kullanılan veri seti tanıtılmakta ve yüz tespiti yöntemlerinin başarımlarını artırmak amacıyla iki yeni sınırlayıcı kutu belirleme yöntemi önerilmektedir.

Üçüncü bölümde; sayısal değerlendirme metrikleri tanıtılmakta ve bu metrikler kullanılarak yüz tespiti yöntemleri ve sınırlayıcı kutu belirleme yöntemlerinin başarımları sayısal olarak analiz edilmektedir. Ayrıca, yöntemlerin başarımları görsel olarak değerlendirilmektedir.

Dördüncü bölümde ise; tez çalışmasında elde edilen sonuçlar değerlendirilmekte, önerilerde bulunulmakta ve gelecekte yapılması hedeflenen çalışmalardan bahsedilmektedir.

1. BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

Yüz tespiti, bir görüntüdeki yüzleri algılamak ve konumlarını belirlemek için kullanılan bir işlemdir. Bu işleme ilişkin detaylar Bölüm 1.1'de verilmiştir. Yüz tespiti, oldukça geniş bir kullanım alanına sahiptir. Bu sebeple, aydınlatma durumundan bağımsız bir şekilde gerçekleştirilmesi büyük önem arz etmektedir. Yetersiz aydınlatma şartlarında gerçekleştirilen yüz tespiti Bölüm 1.1.1'de yer almaktadır. Tez çalışmasının asıl odağı olan yakın kızılötesi görüntüler ve bu görüntülerde yüz tespiti ise Bölüm 1.1.2'de bulunmaktadır.

Literatürde yüz tespitine yönelik oldukça başarılı sonuçlar veren yöntemler bulunmaktadır. Ancak bu yöntemlerin yakın kızılötesi görüntüler üzerinde yüz tespiti başarımının araştırılmasına yönelik bir çalışma bulunmamaktadır. Bu tez çalışmasında, literatürde yaygın olarak kullanılan Dlib [4], FaceBoxes [5], DSFD [6], BlazeFace [7], CenterFace [8], RetinaFace [9], SCRFD [10] ve YuNet [11] yüz tespiti yöntemlerinin yakın kızılötesi görüntüler üzerinde başarımlarının belirlenmesi amacıyla karşılaştırmalı analizleri yapılmıştır. Kullanılan yüz tespiti yöntemleri, Bölüm 1.2'de detaylı olarak verilmiştir.

Son yıllarda yapılan çalışmalarda yüz tespiti yöntemlerinin başarımlarının artırılması için sınırlayıcı kutu belirleme yöntemleri önerilmektedir. Bu yöntemlerin amacı, farklı yüz tespiti modellerinin ürettiği sınırlayıcı kutu topluluklarını kullanarak verilen görüntüdeki yüzleri daha başarılı bir şekilde sınırlayan yeni sınırlayıcı kutuların üretilmesidir. Bu amaç doğrultusunda çalışan, Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS) [12], Ağırlıklı Kutular Birleşimi (Weighted Boxes Fusion, WBF) [13] ve Maksimum Olmayan Ağırlıklandırma (Non-Maximum Weighting, NMW) [14]

yöntemleri literatürde kullanılan yöntemlerdendir. İlaveten; aritmetik ortalama, en temel istatistik hesaplama ölçüsüdür. Buna dayanarak, bahsedilen bu dört yöntem Bölüm 1.3'te açıklanmıştır.

1.1. Yüz Tespiti

Yüz tespiti, görüntü işleme ve makine öğrenmesi tekniklerini kullanarak bir görüntüdeki yüzleri algılama ve tanıma işlemidir. Bu işlem, güvenlik sistemlerinde, otomatik yüz tanıma sistemlerinde, mobil cihazlarda ve sanal gerçeklik uygulamaları gibi birçok alanda kullanılmaktadır [1].

Yüz tespiti, bir görüntüdeki yüzleri algılamak için öncelikle görüntü üzerinde piksel analizleri yapmaktadır. Bu analiz; gözlerin konumu, burun ve ağız gibi yüz üzerinde yer alan organların yerleşimi ve yüzün olası diğer özelliklerini belirlemek için kullanılmaktadır. Daha sonra, buradan elde edilen bilgilerin de ışığında yüz bölgesinin sınırları belirlemekte ve yüzler diğer nesnelerden ayrılmaktadır [15]. Bu yüz bölgesinin tespiti ve diğer nesnelerden ayrılması sürecinde, genellikle makine öğrenmesi algoritmaları ve geliştirilen yapay sinir ağı modelleri kullanılmaktadır. Yüz tespiti için kullanılan bu yaklaşımlar genellikle derin öğrenme tekniklerine dayanmaktadır. Derin öğrenme, büyük veri setleri üzerinde eğitilen ve karmaşık ilişkileri öğrenen yapay sinir ağlarını kullanan yöntemlerdir. Bu yöntem yüzlerin; farklı açılarda, farklı kafa pozisyonlardan ve farklı aydınlatma koşullarında nasıl göründüğünü öğrenmek ve bu bilgiyi kullanarak yüzleri doğru bir şekilde tespit etmek konusunda oldukça başarılı sonuçlar vermektedir. İlaveten, bu yöntemlerin kullanılması ile gerçekleştirilen yüz tespiti, bir görüntüdeki yüzleri algılamakla kalmaz, aynı zamanda yüzlerin özelliklerini analiz edebilir. Örneğin, yüz ifadelerini tanıyabilir, cinsiyeti tahmin edebilir veya yaş tahmini yapabilir. Bu özellikleri sayesinde yüz tespiti; güvenlik, sağlık ve endüstri gibi birçok alanda kullanılabilecek niteliktedir [1].

1.1.1. Yetersiz Aydınlatma Şartlarında Yüz Tespiti

Günümüzde yüz tespiti yöntemleri; güvenlik sahasında, sınır kontrolünde, mobil cihazlarda, fotoğrafçılıkta, sağlık sektöründe ve insan-bilgisayar etkileşiminin olduğu

daha birçok alanda önemli rol oynamaktadır [1]. Bahsi geçen alanlarda gerçekleştirilen yüz tespitinin, uygun aydınlatma şartlarında yapılabildiği gibi yetersiz aydınlatma şartlarında da yapılabilmesi gerekmektedir. Örneğin, bir sınır noktasında geçiş kontrolü sırasında veya seyir halindeki araç içinde hem gündüz hem de gece vakitlerinde yüz tespiti yapılabilmesi kritik önem arz etmektedir. Literatürde yetersiz aydınlatma şartları altında yüz tespiti yapılmasına yönelik oldukça az sayıda çalışma bulunmaktadır. Bu çalışmalar genellikle, yüz tespiti aşaması öncesinde çeşitli görüntü işleme operatörleri kullanılarak girdi görüntünün iyileştirmesine yönelik yaklaşımlardır.

Cho vd. çalışmalarında [16], yetersiz aydınlatma şartlarında yüz tespiti yapılması zorluğuna çözüm aramaktadır. Gece görüşlü kamera ile elde edilmiş görüntülere odaklanılmış çalışmada, karanlık görüntüleri daha aydınlık ve anlaşılabilir hale getirmek üzere görüntü işleme operatörleri uygulanmaktadır. Daha aydınlık hale getirilen görüntü içerisinde, öncelikle hedef kişinin vücudu tespit edilmekte, ardından, tespit edilen kişinin üst yarısı çıkarılmaktadır. Çıkarılan bölge üzerine, omurga ağ olarak VGG Net-16 modelinin kullanıldığı ve R-CNN yaklaşımının benimsendiği bir yöntem ile yüz tespiti yapılmaktadır.

Li vd. [17], gece vakitleri gibi yetersiz aydınlatma koşullarında yapılan yüz tespitinin zorluğuna değinerek, gerçek zamanlı çalışabilecek bir yüz algılama yaklaşımı önermiştir. Önerilen çalışmada, öncelikle girdi karanlık görüntünün kontrastını artırmak için histogram eşitleme (histogram equalization) uygulanmıştır. İkinci aşama olarak, iki piksel değeri arasındaki farkların karşılaştırıldığı İleri Derece Normalleştirilmiş Piksel Farkı (Promotion Normalized Pixel Difference, PRO-NPD) [18] yöntemi kullanılmıştır. Son olarak ise, PRO-NPD neticesinde elde edilen değerin en uygun alt kümesini ve kombinasyonlarını öğrenmek için ikinci dereceden derin ağaç (deep quadratic tree) yaklaşımı benimsenmiştir. Böylece, girdi olarak verilen karanlık gece görüntülerinin iyileştirilmesi ve elde edilen görüntüler üzerinde yüz tespitinin yapılması gaye edinilmiştir.

Wang vd. [19] tarafından önerilen Yüksek-Düşük Adaptasyon (High-Low Adaptation, HLA) çalışmasında, karanlık (düşük değerli) piksel değerlerinin artırılması, aydınlık (yüksek değerli) piksel değerlerinin ise azaltılması yaklaşım benimsenmektedir. Çift

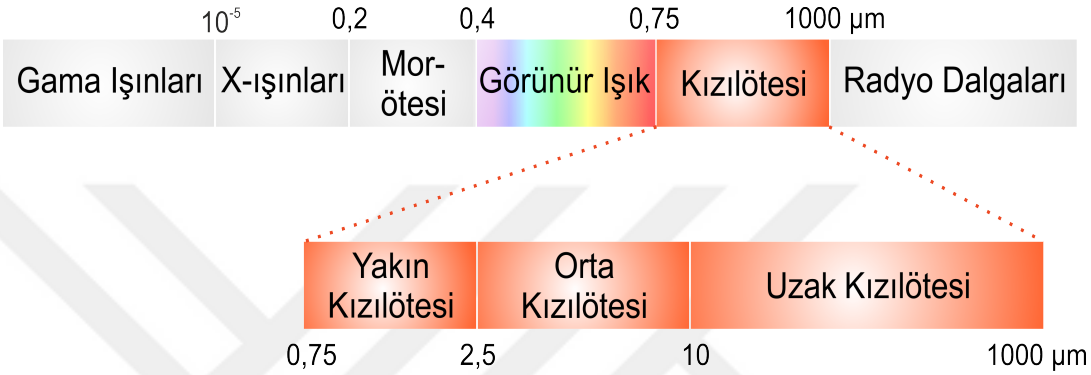
yönlü olarak çalışan bu yaklaşım, her iki durumun da birbirine yaklaşarak, yüz tespiti için daha verimli olan aydınlatma değerlerine yaklaşmasını sağlamaya çalışmaktadır. Geliştirilen yapı kapsamındaki test ve karşılaştırma süreçlerinde Çift Çekim Yüz Tespit Yöntemi (Dual Shot Face Detector, DSFD) [6] kullanılmıştır.

Bahsi geçen çalışmalarda görüleceği üzere, yetersiz aydınlatma durumlarında verimli bir şekilde yüz tespiti yapılamaması problemine yönelik çalışmalar genel olarak girdi görüntülerin iyileştirilmesi prensibine dayanmaktadır. Ancak kullanılan görüntü iyileştirme yöntemleri, ilave işlem gücü ihtiyacı ve işlem süresi getirmektedir. Ayrıca, araç içi gibi kullanıcının görüntüsünün gerçek zamanlı olarak alınması gereken durumlarda, gece görüşlü kamera kullanıcıyı rahatsız edici niteliktedir. Bu sebeple, yetersiz aydınlatma şartlarında yüz tespiti yapılabilmesini sağlamak amacıyla yakın kızılötesi kameralar kullanılabilmektedir.

1.1.2. Yakın Kızılötesi Görüntülerde Yüz Tespiti

Kızılötesi ışık, 1800 yılında gök bilimci William Herschel tarafından keşfedilmiştir. Herschel üçgen bir prizma kullanarak güneşten gelen ışığın kırılmasını sağlamıştır. Bu süreçte, oluşan tayfın tüm renklerinin altına termometre yerleştirmiştir. Termometreleri incelediğinde, en çok ısınan termometrenin kırmızının altındaki olduğunu ve kırmızı renge gidildikçe yaşanan sıcaklık artışını görmüştür. Bu durum gök bilimciyi, kırmızının da ötesinde bir şeyler mi var düşüncesine sürüklemiştir. Böylelikle, yapılan ardışık deneyler sonucunda kızılötesi bölgenin varlığı keşfedilmiştir. Kızılötesi bölge; dalga boyları görünür ışıktan daha uzun olup yaklaşık 750 nanometreden mikrodalgalara kadar olan kesim olarak adlandırılmıştır. Kızılötesi bölge; yakın kızılötesi, orta kızılötesi, uzak kızılötesi olmak üzere alt gruplara ayrılmaktadır. Bunlar içerisinde; yakın kızılötesi, insan gözünün görebildiği dalga boyu aralığına en yakın kızılötesi ışındır. Yakın kızılötesi, görünür ışığa benzer optik kurallara sahiptir. Bu nedende, görünür ışığın yeterli olmadığı durumlarda görüntü oluşturmak için kullanılmaktadır. Ancak yakın kızılötesi ışınlar, görünür ışık dalga boyu aralığının dışında yer aldığından bu ışınları çıplak gözle görmek mümkün değildir [20].

Şekil 1.1'de yer alan ışık tayfında (elektromanyetik spektrum) görüleceği üzere, insan gözünün görebildiği dalga boyu aralığı (0,4 µm - 0,75 µm) ışık tayfında çok küçük bir kısmına karşılık gelmektedir [21]. Bu sebeple, gözle görülemeyecek görüntülere de erişim sağlamak üzere farklı ışıma türleri geliştirilmiştir. Bunlar içerisinde, görünür ışığa en yakın olan yakın kızılötesi ışınlardır. Yakın kızılötesi ışınlar, kızılötesi ışınların 0,75 µm ile 2,5 µm dalga boyları aralığını kapsamaktadır.



Şekil 1.1. Işık tayfı

Yakın kızılötesi ışınlar, genellikle iletişim cihazları ve kamera vb. gece görüş sistemlerinde kullanılmaktadır. Bu ışınlar, görüntüleme sistemlerini engelleyen pus, sis ve duman gibi farklı durumlarda da görüntü alabilmektedir. Bu nedenle, özellikle normal olmayan durumlarda görünür ışıkla görülebilenden daha fazla bilgi ihtiva eden bir görüntü elde edilmesi mümkün olabilmektedir. Diğer taraftan; yakın kızılötesi kameralar, insanın kendi yaydığı radyasyonu algılayabilme kabiliyetine sahiptir. Bu sayede, aydınlatma durumundan bağımsız olarak çalışabilmektedirler. Bu durum, gece görüş sistemlerinin hem askeri hem de sivil amaçlı kullanımlarına imkan tanımaktadır. Gece görüş sistemlerine ilaveten; tarım ürünlerinin kalitelerinin saptanması, kimyasalların ayırt edilmesi, bazı göz rahatsızlıklarına tanı konması ve astronomi gibi çok farklı alanlarda da yakın kızılötesi kameralar kullanılmaktadır. Kullanılan bu kameraların birçoğu, geceleri görüntüleme yapabilmek için menzili genellikle 500 metreyi aşmayan kızılötesi ışık yayan diyotlara (LED) sahiptir. Bu diyotların dalga boyu, yakın kızılötesi alt kümesinde yer almaktadır [22].

Yakın kızılötesi, hem sivil hem de askeri pek çok uygulamada yaygın olarak kullanılmaktadır. Sivil kullanımı oldukça yaygın olduğundan hemen hemen her türlü

ortamda yakın kızılötesi ışığa maruz kalınmaktadır. Yakın kızılötesi ışığa maruz kalınması gözler ve vücut için problem oluşturmayacak niteliktedir. Ayrıca kullanıcısını rahatsız etmiyor olması sebebiyle, yakın kızılötesi görüntüler bir çok alanda olduğu gibi yüz tespiti için de kullanılmaktadır. Bunun en yakın örneği, akıllı telefonlardaki yakın kızılötesi ışığı temel alan yüz tanıma modülleridir [23].

Dash vd. [24]; farklı kafa pozisyonu, arka plan ve aydınlatma koşulu gibi yakın kızılötesi görüntülerde yüz tespit yöntemlerinin başarımını etkileyebilecek etkenleri araştıran bir çalışma gerçekleştirmiştir. Bu çalışmada, yakın kızılötesi görüntülerde yüz tespiti için bahsi geçen etkenlerin olumsuz etkilerini azaltacak yeni bir yüz tespiti yöntemi önermişlerdir. Önerilen yöntemde; aydınlatma koşulları ve algılayıcı kaynaklı yetersiz aydınlatma problemiyle başa çıkmak için görüntü iyileştirme teknikleri ve Hızlı Ardışık Ortalama Niceleme Dönüşümü (Fast Successive Mean Quantization Transform, FastSMQT) kullanılmıştır. Daha sonra, orijinal Seyrek Eleme Ağı (Sparse Network of Winnows, SNoW) sınıflandırıcısından esinlenilerek yeni bir SNoW yaklaşımı geliştirilmiştir. Geliştirilen yaklaşımın amacı, mevcut sınıflandırıcının güncellenmesi ile çalışma hızını artırmak olmuştur. Son olarak ise, FastSMQT ve güncellenmiş SNoW yöntemlerinden elde edilen çıktılar beraber analiz edilerek yakın kızılötesi görüntüde yüz tespiti yapılmıştır. Bu yaklaşımla yapılan yüz tespitinin, çok sayıda insan bulunan kalabalık ortamlarda ve kafa pozlarının farklı olduğu durumlarda başarılı olabileceği öne sürülmüştür.

Gao vd. [25], yetersiz aydınlatma koşullarında görünür spektruma dayalı olarak yüz bölgelerinin tespit edilmesi zorluğuna karşılık, yakın kızılötesi görüntüler için bir yüz tespiti yöntemi önermiştir. Önerilen yöntem, Performans Bozulma Farkı (Performance Degradation Difference, PDD) modeline dayalı olarak çalışan ve farklı aydınlatma koşullarına uyarlanabilir çok modlu bir görüntü birleştirme stratejisine sahiptir. Önerilen stratejideki PDD modeli, ışık yoğunluğunun algılama performansı üzerindeki etkisinin modellenmesiyle elde edilmiştir. Ayrıca, yetersiz aydınlatma koşullarında yüz tespiti performansını daha da artırmak için yüz bölgesinin merkez noktası tahminine yönelik bir yapay sinir ağı tanıtılmıştır. Merkez noktaya dayalı çalışan bu yöntemin, çapa kutusunun (anchor box) neden olduğu sorunları önleyebilecek nitelikte olduğu öne sürülmüştür.

1.2. Yüz Tespit Yöntemleri

Gerçekleştirilen tez çalışmasının ana odağı olan yüz tespiti yöntemlerinin temeli nesne tespit yöntemlerine dayanmaktadır. Nesne tespit yöntemleri, son yirmi yılda devrim niteliğinde büyük bir ilerleme göstermiştir. Bu ilerlemenin derin öğrenme yöntemlerinin kullanılmasından kaynaklandığı bilinmektedir [3]. Bu sebeple; bu çalışmada derin sinir ağları ile yüz öğrenimini gerçekleştiren ve başarısı kanıtlanmış güncel yüz tespit yöntemleri kullanılmıştır.

Derin sinir ağı modelleri, yüz tespiti için kullanılan etkili yöntemlerin başında gelmektedir. Bu modeller, yapay sinir ağlarını kullanarak yüzleri algılamak, konumlandırmak, tanımak ve sınıflandırmak gibi görevleri üstlenmektedirler. Derin sinir ağı modelleri ile geliştirilmiş yüz tespiti yöntemleri, farklı prensiplere ve algoritmalara dayanarak yüz tespiti gerçekleştirmektedirler. Bu sebeple, bazıları daha hızlı çalışırken bazıları daha hassas veya daha yüksek doğruluk oranlarına sahip sonuçlar verebilmektedir. Bu bölümde, literatürdeki güncel ve popüler yüz tespiti yöntemlerine alt başlıklarda yer verilmiştir.

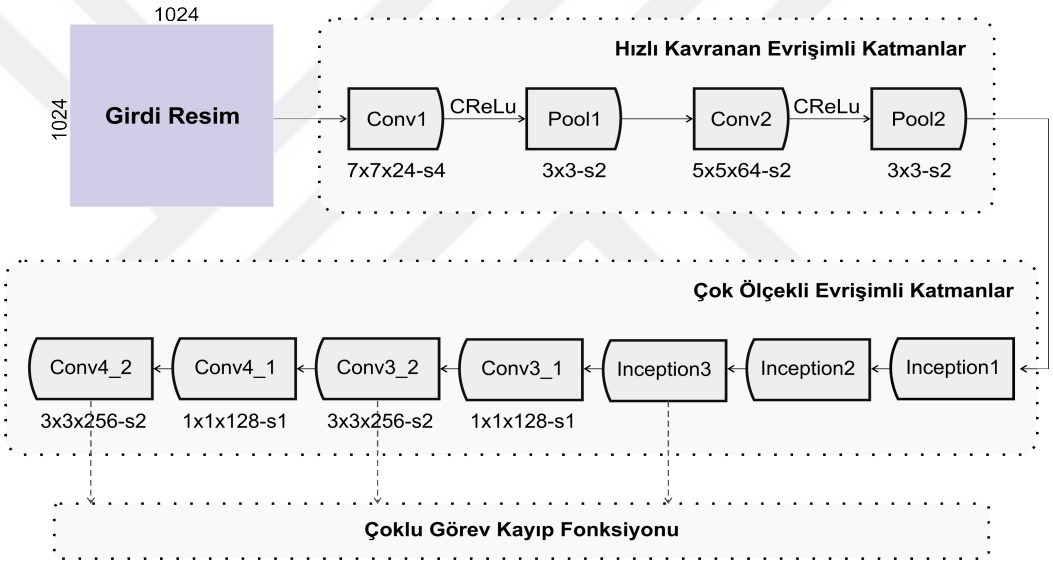
1.2.1. Dlib

Dlib [4], C++ programlama dilinde yazılmış açık kaynaklı bir makine öğrenmesi kütüphanesidir. Farklı alanlarda ve projelerde sıklıkla kullanılan bu kütüphane, makine öğrenme algoritmaları, görüntü işleme yöntemleri ve renk uzayı dönüşümleri, sayısal hesaplama fonksiyonları gibi birçok fonksiyonu da bünyesinde barındırmaktadır.

Dlib; OS X, MS Windows, Linux, Solaris, BSD ve HP-UX işletim sistemlerini desteklemektedir. İlaveten; Android, IOS gibi mobil platformlarda ve Raspberry Pi, Nvidia Jetson, Tinkerboard gibi gömülü donanımlar üzerinde de çalışabilmektedir. Bu geniş çalışma yelpazesi; Dlib'in, cep telefonları, tabletler, gömülü cihazlar, robotik ve yüksek performanslı bilgi işlem ortamları başta olmak üzere hem akademi hem de endüstride popüler olarak kullanılmasını sağlamaktadır. Dlib'in içerisinde; obje takibi, obje poz tahmini ve özellikte yüz tespiti için kullanılan güçlü araçların olması ve bunları kolay kullanılıp sistemlere kolay entegre edilebiliyor olması tercih edilme sebepleridir.

1.2.2. FaceBoxes

FaceBoxes [5], yalnızca tek bir evrişimli sinir ağı içeren ve uçtan uca eğitilebilen bir yüz tespit yöntemidir. Bu yüz tespit yöntemi; Şekil 1.2'de mimarisine yer verildiği üzere, Hızlı Kavranan Evrişimli Katmanlar (Rapidly Digested Convolutional Layers, RDCL) ve Çok Ölçekli Evrişimli Katmanlar (Multiple Scale Convolutional Layers, MSCL) olarak adlandırılmış katmanlardan oluşan bir ağ yapısına sahiptir. Ayrıca bu ağ yapısı, görüntü üzerinde aynı yoğunluğa sahip farklı tespit noktaları üretme stratejisini de benimsemektedir. Bu strateji, küçük boyutlu yüzlerin de tespit edilebilme oranını artırmaktadır.



Şekil 1.2. FaceBoxes mimarisi
[5]

RDCL, ilgili yüz tespit yönteminin gerçek zamanlı çalışma hızına ulaşmasını sağlamak için tasarlanan katmanları barındırmaktadır. Bu katmanlar, çıkış kanallarının sayısını azaltarak, giriş uzamsal boyutunu seçilecek uygun çekirdek boyutuna göre hızlı bir şekilde küçültmek için tasarlanmıştır.

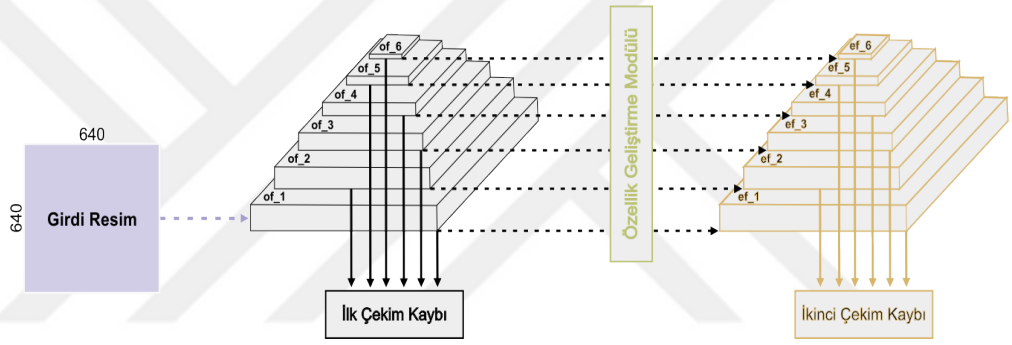
MSCL, çok kategorili nesne algılama yaklaşımlarında sınıftan bağımsız bir önerici olarak geliştirilen Bölge Öneri Ağına (Region Proposal Network, RPN) [26] dayalı olarak tasarlanmıştır. Bu tasarım sürecinde; FaceBoxes geliştiricileri tarafından RPN'nin eksiklikleri, zayıflıkları ve sorunları detaylı bir analiz neticesinde tespit edilmiştir. Tespit edilen bu durumlar, ağ yapısı içerisinde hem derinlik hem de genişlik esas

alınarak ağ boyutu boyunca çok ölçekli tasarıma göre geliştirilmek suretiyle sorunlara bir çözüm önerisi getirilmiştir. Ayrıca MSCL, çeşitli ölçeklerdeki yüzleri işlemek için alıcı alanları (receptive fields) zenginleştirmeyi ve farklı tespit noktalarını ayrıştırmayı amaçlamaktadır.

1.2.3. DSFD

Çift Çekim Yüz Tespit Yöntemi (Dual Shot Face Detector, DSFD) [6]; Tek Çekim Algılayıcısı (Single Shot Detector, SSD) [27] mimarisi üzerine üç yeni değişiklik önermektedir. Bu önerilerden ilki; sınıflandırmaların ve regresyonların yürütüldüğü özellik haritalarının hesaplaması üzerine yapılmıştır. Yapılan çalışmada SSD ağından farklılaşan temel unsur, çeşitli derinliklerdeki altı özellik haritasının "geliştirilmiş" özellik haritasına dönüştürülmesidir. Bu dönüştürme görevi için, Özellik Geliştirme Modülü (ÖGM) adı verilen bir modül tasarlanmıştır. Bu modülün amacı, nesne sınıflandırıcılarını ve sınırlayıcı kutuları daha esnek ve güçlü özelliklerle besleyebilmektir. Önerilerden ikincisi; ağ eğitimi sırasında tespit kaybını en aza indirmeyi hedefleyen bir kayıp fonksiyonu yaklaşımıdır. Bu yaklaşımda; geleneksel nesne tespit kayıp fonksiyonlarından farklı olarak, ağda yalnızca farklı seviyeler için değil, aynı zamanda çekim olarak tanımlanmış iki farklı yapı için de aşamalı kayıp hesabı uygulanmaktadır. Çift Çekim Yüz Tespit Yöntemi ismi; bu çift akışlı tasarım doğrultusunda süreçlerin ilerliyor olması sebebiyle ilgili algılayıcıya verilmiştir. Yüz tespit yönteminin getirdiği üçüncü öneri ise; gerçekleştirilen tahminlerin resim üzerindeki yüzler ile daha etkili bir şekilde eşleştirilmesini amaçlayan bir stratejidir. Bu stratejide, yüz tespiti sırasında yüz bölgesi alanının net bir şekilde belirlenememesi sorununa odaklanılmıştır. Bu sorun, eğitimde tespit edilen pozitif ve negatif sınırlayıcı kutuların sayıları arasında negatif yönde bir dengesizliğe yol açarak, daha yavaş ve daha az kararlı süreç optimizasyonuna neden olmaktadır. Bu nedenle, odaklanılan sorunu çözmek için veri arttırma stratejisi benimsenmiştir. Bu bağlamda, eğitim sırasında giriş görüntülerine %40 oranında veri arttırma (data augmentation) uygulanmıştır. Uygulanan bu işlemin, pozitif sınırlayıcı kutuların sayısını arttırarak pozitif-negatif sınırlayıcı kutu dengesinin sağlanması ve daha kapsayıcı yüz bölgesi tespitlerinin yapılmasına olanak sağladığı öne sürülmüştür.

DSFD; PyramidBox [28] ve S3FD [29] ile aynı genişletilmiş VGG16 omurga ağını kullanmaktadır. Bu ağ yapısında, sınıflandırma katmanlarından önce bazı katmanlar kesilmiş ve yardımcı yapılar eklenerek ağ yapısı desteklenmiştir. İlgili yapıya ilişkin mimari Şekil 1.3'te yer almaktadır. Ağ yapısı içerisinde, altı farklı evrişim katmanı, ilk çekim algılama katmanları olarak belirlenmiştir. Bu katmanlar, altı adet orijinal özellik haritalarını oluşturmaktadır. Daha sonra, önerilmiş yapı içerisinde tek çekim algılayıcısını çift çekim algılayıcısına genişletmek için kullanılan ÖGM ile bu orijinal özellik haritaları altı gelişmiş özelliğe aktarılmaktadır. Bu özellik haritaları, orijinal olanlarla aynı boyutlara sahiptir ve ikinci çekim algılama katmanlarını oluşturmak için SSD tarzı bir yaklaşımı benimsemişlerdir.



Şekil 1.3. DSFD mimarisi

[6]

Aynı omurga ağını kullanan PyramidBox ve S3FD gibi yüz tespit yöntemlerinden farklı olarak, ÖGM'deki alıcı alan genişlemesinden sonra üç boyutlu adım (stride) uygulanması, tespit noktası ve alıcı alan için eşit orantı aralığı ilkesinin karşılanması bahsedilen yüz tespit yöntemi için gereksiz görülmüştür. İlaveten, orijinal ve gelişmiş çekimlerin, İlk Çekim Kaybı ve İkinci Çekim Kaybı olarak adlandırılan iki farklı kaybı olması sebebiyle, DSFD geliştiricileri yüz tespit yöntemlerinin diğer yöntemlerden daha esnek ve güçlü olduğunu savunmaktadır.

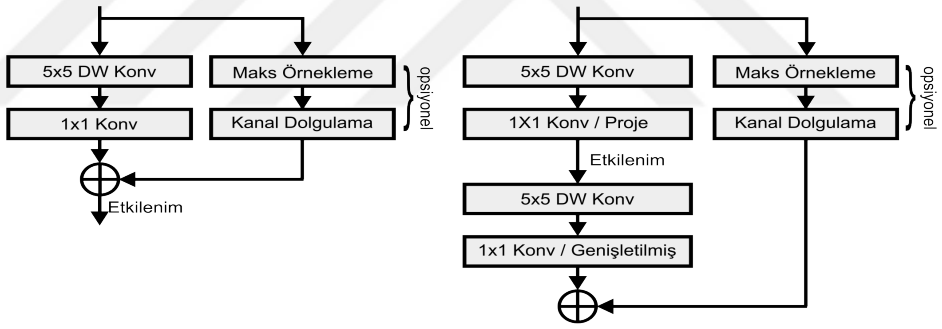
1.2.4. BlazeFace

BlazeFace [7], mobil GPU ortamları için özel olarak tasarlanmış, hafif ve oldukça hızlı çalışan bir yüz tespit yöntemidir. Bahsedilen yüksek hız, algılayıcının geliştiricileri tarafından, amiral gemisi mobil cihazlar için 200-1000+ FPS olarak ön görülmüştür. Bu

güçlü gerçek zamanlı performans; 2B/3B yüz anahtar noktası veya yüz bölgesi geometri tahmini, yüz özellikleri veya mimik sınıflandırması ve yüz bölgesi segmentasyonu gibi yüz tespitlerinin oldukça önemli olduğu problemlere geliştirilecek çözümleri hızlandırarak performanslarını doğrudan etkileyecek kabiliyete sahiptir.

BlazeFace, Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS)'ya [26, 27, 30] alternatif olarak geliştirilmiş bir bağlantı çözme stratejisine ve SSD [27]'den değiştirilmiş GPU dostu bir bağlantı şemasına sahiptir. Bu şema, MobileNetV1/V2 'den [31, 32] esinlenmiştir, ancak onlardan farklı olarak daha düşük ağırlıklara sahip bir özellik çıkarma ağı içermektedir.

BlazeFace mimarisinde; birçok gözlem ve deney yapılarak, popüler olarak kullanılan 3×3 derin evrişim katmanının 5×5 derin evrişim katmanıyla değiştirilmesi önerilmiştir. Bu yaklaşımın, tespit sürecini hızlandırmak için modeli daha sığ hale getirdiği düşünülmektedir (Şekil 1.4, (a)).



Şekil 1.4. BlazeFace mimarisi

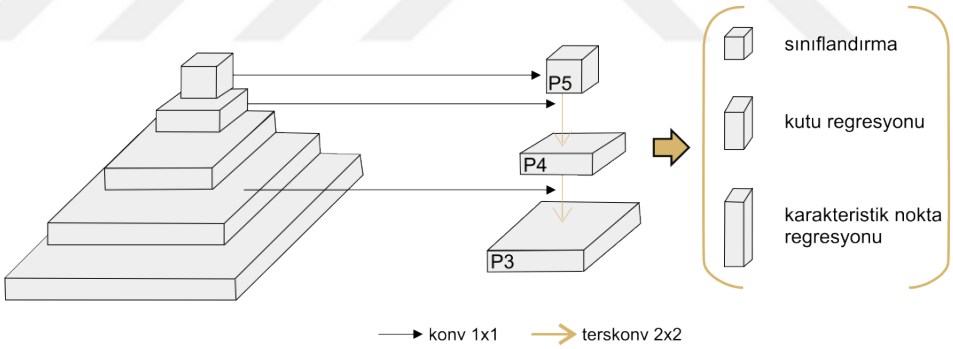
[7]

BlazeFace'in esinlendiği ağlardan biri olan MobileNetV2 [32], doğrusal olmayan bir durumla ayrılmış ve birbirini takip eden derinliği artıran genişleme evrişim katmanları ve derinliği azaltan noktasal evrişim katmanlarına sahiptir. Ancak BlazeFace, ara tensörlerde daha az sayıda kanala uyum sağlamak için, ağ yapısındaki artık bağlantıların genişletilmiş kanal yapısında çalışması için bu aşamaları değiştirmektedir. Bu sayede, derin evrişim katmanının, bu iki evrişim katmanı arasına bir başka katman eklenmesine olanak tanımaktadır. Bu yapı, BlazeFace'in daha yüksek soyutlama düzeyli katmanları için tercih edilen çift BlazeBlock'un özünü oluşturmaktadır (Şekil 1.4).

1.2.5. CenterFace

CenterFace [8]; gerçek zamanlı hız ve yüksek doğruluk oranıyla, kişinin yüzü ve yüz karakteristik noktalarının (face landmarks) konumunu aynı anda tahmin etme yetisine sahip bir yüz tespit yöntemidir. Modelin geliştirilmesinde, özellikle düşük ağırlıklı bir ağ yapısı ve daha basit bir yaklaşım benimsenmiştir. Karmaşık modellere sahip yüz tespit yöntemlerinin dezavantajlarının değerlendirildiği çalışmada, hafif ve basit yaklaşımla beraber ilgili algılayıcının güçlü ve etkili bir yüz tespiti ve hizalama sağlaması üzerinde durulmuştur. Yapılan çalışmalar neticesinde; ilgili yüz tespit yönteminin geliştiricileri, semantik haritalar ile olası mevcut yüz bölgesinin öğrenilmesi ve sınırlayıcı kutunun yer alacağı yüz bölgesi ve bu yüz bölgesi üzerinde bulunacak potansiyel yüz işaretlerinin öğrenilmesi üzerine başarı sağlayan bir yöntem geliştirdiklerini öne sürmüşlerdir.

CenterFace mimarisine Şekil 1.5'te yer verilmiştir. Mimari içerisinde, omurga ağ olarak MobileNetV2 [32] ve boyun ağı olarak Özellik Piramidi Ağı (Feature Pyramid Network, FPN) [33] kullanılmıştır.



Şekil 1.5. CenterFace mimarisi
[8]

FPN; tek bir ölçek girdisinden tek bir özellik piramidi oluşturmak için, yanal bağlantılara sahip olan dikey bir mimari yapısına sahip olacak biçimde ilgili şekilde temsil edilmiştir. Mimari tasarımı doğrultusunda; yüz algılama ve hizalama için piramitteki yalnızca bir katman kullanılmıştır. Bu bağlamda, öncelikle yüz bölgesi girdi resim içerisinde sınıflandırılmaktadır. Ardından, sınırlayıcı kutu ile çerçevelenen yüz bölgesinin merkez noktası baz alınarak sınırlayıcı kutunun boyutu ve yüz bölgesi karakteristik noktaları çıkarılmaktadır. Bahsedilen karakteristik noktalar; gözler, dudak kenarları ve burun ucudur. Bu yapıda, konvolüsyon ve ters konvolüsyonlar ile bağlanmış P3-P5

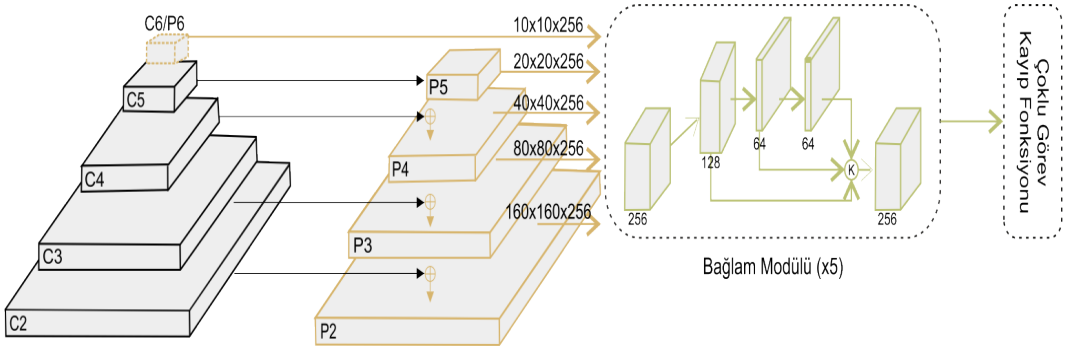
piramitlerindeki değerler ilgili piramidin seviyesini göstermektedir. Burada belirtilen değerler (3-5), girdi görüntünün çözünürlüğünü belirlemede $1/2^{3-5}$ formülasyonu üzerinden kullanılmaktadır.

Yüz tespit sürecinde öncelikle, yüz bölgesi konumunum yer aldığı sınırlayıcı kutu, merkez noktası olarak belirlenmektedir. Ardından sınırlayıcı kutunun boyutu ve yüz işaretleri doğrudan merkez konumdaki görüntü özelliklerine göre elde edilmektedir. Bu durum; yüz algılama ve hizalama işlevini, standart kilit nokta tahmin problemine [34, 35] dönüştürmektedir. Buradan üretilen ısı haritasındaki tepe, yüzün merkezine karşılık gelmektedir. Her zirvedeki farklı görüntü özellikleri ise, yüzün boyutunu ve yüzün kilit noktalarını tahmin etmek için kullanılmaktadır. Bu yaklaşım, literatürdeki algılayıcılara göre hesaplama maliyetini azaltmaya yarayan bir adım olarak sunulmuştur.

1.2.6. RetinaFace

RetinaFace [9]; standart dışı durumlarda verimli bir şekilde yüz bölgesi tespiti gerçekleştirmenin zorluğuna değinerek, bu soruna yönelen bir yüz tespit yöntemi olarak önerilmiştir. Tespit edilen soruna çözüm için geliştirilen yüz tespit yöntemi, ekstra denetimli (extra-supervised) ve öz denetimli (self-supervised) öğrenme tekniklerinin kullanıldığı çoklu görev (multi-task) öğrenimi üzerine kurgulanmıştır. Bu çerçevede, çeşitli yüz ölçeklerinde piksel bazlı yüz bölgesi tespitleri yapabilen güçlü bir yüz tespit yöntemi oluşturulmuştur.

RetinaFace mimarisi Şekil 1.6'da verilmiştir. Mimari incelendiğinde, ilk olarak Özellik Piramidi Ağına (ÖPA), ardından Bağlam Modülüne ve son olarak Çoklu Görev Kayıp Fonksiyonuna sahip bir yapı görülmektedir. ÖPA, 2 boyutlu görüntü üzerinden beş farklı seviyede özellik çıkarmaktadır. Çıkarmış olduğu özelliklerden ilk dördü (P2-P6), önceden eğitilmiş ResNet [36] modelinin artık bloklarının çıkış katmanları (C2-C5) kullanılarak hesaplanmaktadır. Beşinci özellik (C6/P6) ise, 3×3 evrişim katmanı ve iki piksellik adım (stride) uygulanarak elde edilmektedir. Elde edilen tüm bu özellikler, daha fazla bağlamsal bilgiye erişebilmek için beş farklı filtreye sahip bağlam modülüne iletilmektedir. Son olarak, çoklu görev kayıp fonksiyonu ile tespit kaybı hesaplanmaktadır.



Şekil 1.6. RetinaFace mimarisi

[9]

RetinaFace, piksel bazlı yüz lokalizasyonu [37] gerçekleştirmektedir. Bu işlev için üç alt görevi uygulayan bir yapıya sahiptir. Bu görevler; yüz tespiti [38], 2 boyutlu yüz geometrik yapısı algılama [39] ve 3 boyutlu yüz çıkarımıdır [40]. Uyguladığı bu yoğun yüz lokalizasyonu sayesinde yüz tespit yönteminin, farklı açı ve ölçeklerde alınabilecek görüntüler için de daha doğru yüz konumu bilgisi sağladığı öne sürülmektedir. Yapılan yüz tespitlerine ilaveten; göz merkezleri, burun ucu ve ağız köşeleri olmak üzere toplam beş farklı yüz noktası da tespit esnasında çıkarılabilmektedir.

1.2.7. SCRFD

SCRFD [10], düşük çözünürlüklü görüntülerde küçük boyutlu yüzleri algılamanın zorluğuna dikkat çekerek bu soruna odaklanan bir yaklaşım önermiştir. Önerilen yaklaşımda, veri örnekleme (data sampling) ve hesaplama dağılımı (computation distribution) stratejilerinin verimli ve yüksek doğruluk oranlarına sahip yüz tespitleri için oldukça önemli olduğu belirtilmiştir. Bu doğrultuda, iki farklı basit ama etkili yöntem tanıtılmıştır. Bunlardan ilki; kullanılacak veri setinin analizine dayanarak en çok ihtiyaç duyulan noktalarda eğitim örneklerini artıran Örnek Yeniden Dağıtımı (Sample Redistribution, SR) stratejisidir. İkincisi ise; hesaplamayı modelin omurga ağı, boyun ağı ve baş ağı arasında yeniden tahsis eden Hesaplama Yeniden Dağıtımı (Computation Redistribution, CR) stratejisidir. Önerilen SR ve CR stratejileri; birbirini tamamlar niteliktedir. Beraber kullanımlarının başarıyı pozitif yönde doğrudan etkileyeceği düşüncesi ile geliştirmişlerdir.

SCRFD modeli kurgulanırken; hesaplama maliyetini azaltmak için test aşamasında büyük ölçek kullanmak yerine sabit bir VGA çözünürlüğü (640×480) benimsenmiştir. Bu yaklaşımın, yüz algılamayı daha verimli kıldığı düşünülmektedir. Benimsenen bu ölçek ayarı altında, WIDER Face [41] veri setindeki yüzlerin çoğu (%78,93) 32×32 pikselden küçük duruma gelmektedir. Bu nedenle, omurga ağın sığ aşamaları ile tahmin edilebilmişlerdir [10]. Bu sığ aşamalar için daha fazla eğitim örneği elde etmek için SR yöntemi önerilmiştir. Yöntem kapsamında, [1,2, 1,4, 1,6, 1,8, 2,0] ölçekleri rastgele eklenerek veri setindeki küçük boyutlu yüzler genişletilmiştir. Yapılan bu ölçek artırma işleminin yanı sıra, eğitim verileri renk bozulması ve 0,5 olasılıkla rastgele yatay çevirme yöntemleri ile artırılmaktadır.

Yüz tespiti gerçekleştirme aşamasında, bir sınıflandırma modelinin omurga ağını doğrudan kullanmak yetersiz olabilmektedir. Bu nedenle önerilen CR stratejisinde; farklı durumlar altında omurga ağı, boyun ağı ve baş ağı üzerindeki hesaplamayı yeniden tahsis etmek için ağ yapısı araması [42] kullanılmıştır. Bu arama metodu, ResNet [36]'in omurga ağ olduğu RetinaNet [33] üzerine uygulanmıştır. Boyun ağı olarak Yol Toplama Özelliği Piramit Ağı (Path Aggregation Feature Pyramid Network Network, PAFPN) [43] ve baş ağında 3×3 boyutunda evrişim katmanları kullanılmıştır.

Bir yüz tespit yönteminin yapısı; hesaplamanın dağılımını, doğruluğunu ve verimliliğini belirlemede büyük önem taşımaktadır. Bu bilinçle, omurga ağı, boyun ağı ve baş ağı üzerinde farklı konfigürasyonlara sahip model mimarileri üretme çalışması yapılmıştır. İlk adımda, boyun ağı ve baş ağı bileşenlerini sabitlerken hesaplamanın omurga ağı parçaları içinde yeniden tahsisi sağlanmıştır. Sonrasında; elde edilen omurga ağı üzerinde optimize edilmiş hesaplama dağılımına dayanarak, hesaplamanın omurga ağı, boyun ağı ve baş ağı boyunca gerçekleştirilmesi üzerine çalışılmıştır. Her iki aşamada da optimizasyon yapılarak, yüz tespiti için nihai optimize edilmiş ağ tasarımına ulaşılmıştır.

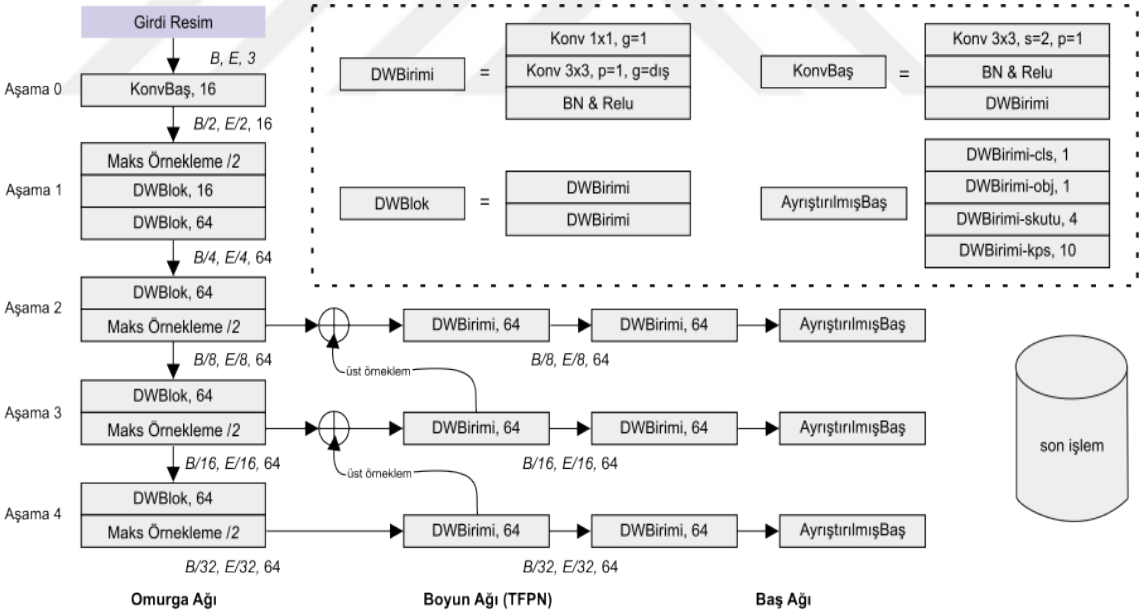
1.2.8. YuNet

Model mimarilerinin ağır olması ve yüksek hesaplama maliyetleri, birçok yüz tespit yönteminin mobil ve gömülü cihazlarda kullanılabilmesini zorlaştırmaktadır. Çünkü bu cihazlarda işlenebilecek model boyutu ve oluşacak gecikme süresini karşılayabilme

yetisi oldukça kısıtlıdır. Bu durum göz önüne alınarak, hedef cihazlar için özel olarak tasarlanmış ve milisaniye düzeyinde çalışabilecek bir yüz tespit yöntemi olan YuNet [11] geliştirilmiştir. YuNet; hafif, hızlı ve yüksek doğruluk oranına sahip bir yüz algılama modelidir. Oldukça hafif bir ağ yapısı üzerinden tasarlanan bu model, tespiti zor olan küçük boyutlu yüzlere odaklanmayı ve tespiti kolay olan büyük boyutlu yüzlerin hesaplama maliyetini olabildiğince azaltmayı ilke edinmiştir.

YuNet, yalnızca 75856 parametreye sahiptir. Çok sınırlı sayıda parametre, çok düşük gecikme süresi ve umut verici doğruluk beklentisi üzerine kurgulanan YuNet, geliştiricileri tarafından doğruluk ve hız arasında en iyi dengeye sahip yüz tespit yöntemi olarak tanımlanmıştır.

YuNet'in mimarisi Şekil 1.7'de verilmiştir. Bir omurga ağı, küçük özellikli bir piramit ağı (A Tiny Feature Pyramid Network, TFPN) bulunduran boyun ağı ve bir baş ağı olmak üzere ilgili mimari üç bölümden oluşmaktadır.



Şekil 1.7. YuNet mimarisi
[11]

Omurga ağı, ağ mimarisinin ana parçasıdır. Bu parça, yüz tespitleri için özniteliklerin çıkarılması için kullanılmaktadır. Geliştirme sürecinde özellikle hafif ve verimli olması için çaba gösterilmiştir. Bu bağlamda, model eğitimi için veri örnekleme (data sampling) stratejisi önerilmiştir. Bu stratejinin; özellikle hafif algılayıcıların tespit doğruluğunu

artırabilir nitelikte olduğu savunulmaktadır. Beş aşama olarak tanıtılmış olunan omurga ağı, MobileNet [31] örnek alınarak oluşturulmuştur.

Boyun ağı, çok ölçekli özellikleri birleştirerek daha yüksek bir özellik düzeyi elde etme görevini üstlenmiştir. Boyun ağının ana parçası olan TFPN, FPN'den [33] ve derinlemesine ayrılabilir evrişimlerden (depthwise separable convolution) faydalanılarak türetilmiştir. FPN'deki evrişim katmanları, ağır bir hesaplama maliyetine sahiptir ve birçok parametrenin eğitilmesini gerektirmektedir. Bu maliyetleri azaltmak amacıyla TFPN boyun ağı, standart evrişim katmanları yerine derinlemesine ayrılabilir evrişimleri kullanmaktadır. Ayrıca, FPN'deki 3×3 evrişim katmanları DWBirimi ile değiştirilmiştir. Yapılan bu değişiklik, parametre sayısını FPN'den %12'ye kadar azaltabilecek ve hesaplama maliyeti de büyük ölçüde düşürebilecek niteliktedir.

Baş ağı içerisinde, sınırlayıcı kutuyu doğrudan tahmin etme prensibi (anchor-free) benimsenmiştir. Bu prensip doğrultusunda çalışan Ge ve ark. çalışmasından [44] ilham alınarak basit optimum taşıma ataması (simOTA) [45] kullanılmıştır. Gerçekleştirilen bu yaklaşımın RetinaFace [9] gibi referans tabanlı sınırlayıcı kutu tahmin etme (anchor-based) yaklaşımlarına göre yaklaşık iki kat hızlı çalıştığı belirtilmiştir.

1.3. Sınırlayıcı Kutu Belirleme Yöntemleri

Sınırlayıcı kutu topluluğundan yeni bir sınırlayıcı kutu belirleme yöntemi (bounding box ensemble), birden fazla sınırlayıcı kutu tahminini bir araya getirip, onların birlikte analiz edilmesi ile daha başarılı ve daha güvenilir bir sınırlayıcı kutu elde etmeyi amaçlayan bir yöntemdir [46]. Bu sayede, farklı yaklaşımların avantajları bir araya getirilerek; daha yüksek doğruluk oranına sahip, daha hassas ve daha güvenilir sonuçlar elde edilebilmektedir.

Nesne tespiti yöntemlerinin en sık kullandığı yaklaşımlardan biri, sınırlayıcı kutu ile tespit edilen nesne konumunun belirlenmesidir. Sınırlayıcı kutular, bir nesnenin koordinatlarını (sol üst ve sağ alt köşelerin piksel konumları) içeren dört değeri temsil etmektedir. Ancak tek bir sınırlayıcı kutu tahmini, bazen yanıltıcı veya hatalı olabilmektedir. Örneğin, tespiti hedeflenen nesnenin bir kısmı sınırlayıcı kutu içerisine dahil edilememiş veya yanlış bir bölge dahil edilmiş olabilmektedir. Bu tür sorunları çözmek için sınırlayıcı kutu belirleme

yöntemleri kullanılmaktadır. Bu yöntemlerde, farklı model veya algoritmaların ürettiği farklı sınırlayıcı kutu tahminlerinin beraber analizi sağlanmaktadır. Bu sayede, farklı yaklaşımların avantajları bir araya getirilerek; daha yüksek doğruluk oranına sahip, daha hassas ve daha güvenilir sonuçlar elde edilebilmektedir.

Topluluktan yeni bir sonuç çıkarma prensibi, her bir modelin veya algoritmanın ürettiği sonuçları birleştirerek daha kapsamlı ve doğru bir sonuç elde etmeyi hedeflemektedir. Bu durum, farklı modellerin veya algoritmaların güçlü yanlarını birleştirerek zayıf yönlerinin telafi edilmesini sağlamaktadır. Bu yaklaşım üzerine kurgulanan sınırlayıcı kutu belirleme yöntemleri, tek bir sınırlayıcı kutudan alınan performansın artırılması için çalışmaktadır. Bu yöntemler; ölçeklendirme, pozisyon ve perspektif değişiklikleri gibi çeşitli zorluklarla karşılaşılan karmaşık durumlarda etkili olabilmektedir. Ayrıca, farklı veri kaynaklarından elde edilen tahminleri birleştirmek için de kullanılabilmektedir. Özellikle nesne tespiti ve algılanması gibi uygulamalarda, bu yöntem daha başarılı ve istikrarlı sonuçlar sağlayabilmektedir.

Bu bölümde; Maksimum Olmayan Bastırma, Maksimum Olmayan Ağırlıklandırma, Ağırlıklı Kutular Birleşimi ve Ortalama Alma yöntemleri alt başlıklarda yer almaktadır.

1.3.1. Maksimum Olmayan Bastırma

Maksimum olmayan bastırma (Non-Maximum Suppression, NMS) [12], bir yüz/nesne detektörü tarafından oluşturulmuş örtüşen sınırlayıcı kutuları kaldırmak için kullanılan bir son işleme tekniğidir. Yüz algılama, nesne takibi, görüntü ve videolarda nesne algılama gibi birçok bilgisayar görü alanındaki farklı görevlerde yaygın olarak kullanılmaktadır.

NMS'nin arkasındaki temel fikir, en yüksek güven puanına (confidence score) sahip sınırlayıcı kutuyu seçmektir. Bu seçim süreci, en yüksek puana sahip olan sınırlayıcı kutu ile kullanıcı tarafından belirlenen eşik değerinden daha yüksek oranda örtüşen diğer sınırlayıcı kutuları kaldırmak suretiyle ilerletilmektedir. Bu işlem, tüm sınırlayıcı kutular için ayrı ayrı uygulanmaktadır. Böylelikle aynı nesneyi temsil etme potansiyeli olan sınırlayıcı kutular arasından en yüksek puana sahip olan dışındakiler elenmektedir.

Genellikle nesne algılayıcıları tarafından aynı nesne için çok sayıda sınırlayıcı kutu üretilmektedir. Bu durum; doğru pozitiflerin tespitini sağladığı gibi yanlış pozitiflerin de ortaya çıkmasına sebebiyet vermektedir. NMS, yalnızca tespit edilen nesnelere karşılık gelen en ilgili sınırlayıcı kutuları seçerek bu yanlış pozitifleri ortadan kaldırma görevini üstlenmektedir. Yanlış veya gereksiz tespitleri ortadan kaldıran bu yöntem, nesne tespit algoritmalarının hesaplama karmaşıklığını azaltmaya da yardımcı olmaktadır.

1.3.2. Maksimum Olmayan Ağırlıklandırma

Maksimum olmayan ağırlıklandırma (Non-Maximum Weighting, NMW) [14], nesne tespiti algoritmalarının doğruluğunu ve hassasiyetini artırmak için kullanılan bir sınırlayıcı kutu belirleme yöntemidir. Bu yöntem, birden fazla nesnenin bir görüntüde aynı bölgede bulunması durumunda doğru sınırlayıcı kutuların seçilmesine yardımcı olmaktadır. NMW, yine aynı görevlerde kullanılan Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS) tekniğinden esinlenilerek geliştirilmiştir. NMS, gereksiz sınırlayıcı kutu algılamalarını filtrelemek ve yalnızca en güvenilir ve örtüşmeyen tahminleri tutmak için kullanılmaktadır. Standart bir NMS yaklaşımında, yüksek güven puanlarına sahip sınırlayıcı kutulara öncelik verilerek belirli bir eşiğin üzerindeki örtüşen kutular elimine edilmektedir (bastırılmaktadır). Bu durum, aynı nesne üzerinde yinelenen veya oldukça benzer konumlandırılan algılamaları ortadan kaldırmaya yöneliktir. Bu yaklaşım, belirlenecek son sınırlayıcı kutunun iyileştirilmesini sağlamaktadır. Yine aynı amaç üzerine kurgulanan NMW'de ise; gerçekleştirilen bu bastırma işlemi sırasında ek ağırlıklandırma faktörleri dahil edilerek elde edilecek çıktı performansının yükseltilmesi için çalışılmaktadır. Bu faktörler; hangi sınırlayıcı kutuların saklanacağını veya bastırılacağını belirlerken, güven puanları, nesne boyutları veya diğer olası niteliklerin dikkate alınmasını içerebilmektedir.

Nesne tespiti algoritmaları çıktılarında; benzer özelliklere sahip nesnelerin çakışması, birbirini kapatması veya sınırlayıcı kutuların alakasız konumlandırılması gibi sorunlar ortaya çıkabilmektedir. NMW, bu sorunları çözmek için kullanılan bir yaklaşımdır. Bu yaklaşımda, görüntüdeki potansiyel nesnelerden oluşan bir dizi sınırlayıcı kutu bulunarak bu sınırlayıcı kutular arasındaki örtüşme değerlendirilmektedir. Değerlendirme sürecinde; öncelikle görüntü üzerinde tespit edilen nesnelerin sınırlayıcı kutuları, bu kutuların

güven puanları ve konumları belirlenmektedir. Ardından, güven puanı en yüksek olan sınırlayıcı kutu, diğer yakın kutularla benzerlik ölçütleri kullanılarak karşılaştırılmaktadır. Benzerlik ölçütleri, genellikle örtüşme alanı, Jaccard veya IoU (Intersection over Union) olarak adlandırılan performans değerlendirme metriği kullanılarak hesaplanmaktadır. Hesaplama sonrasında, her bir nesne için ağırlığı en yüksek olan sınırlayıcı kutu seçilerek diğer yakın olan kutular kaldırılmaktadır. Bu seçme ve kaldırma süreci, bir döngü içinde tüm sınırlayıcı kutular üzerinde tekrarlanmaktadır. Sonuç olarak, her bir nesne için en belirgin ve ağırlığı en yüksek sınırlayıcı kutu elde edilmektedir. Diğer taraftan, bu sınırlayıcı kutular değerlendirilirken her bir sınırlayıcı kutu için bir ağırlık değeri hesaplamaktadır. Bu ağırlıklar, kutunun içerdiği nesnenin önem düzeyine bağlı olarak atanmaktadır. Örneğin, daha büyük veya daha belirgin nesneler daha yüksek bir ağırlık alabilmektedirler. Ardından, sınırlayıcı kutuların örtüşme dereceleri ve ağırlıkları kullanılarak bir puan hesaplanmaktadır. Bu puan, bir kutunun diğerlerine göre ne kadar öncelikli olduğunu temsil etmektedir. Puanlama işlemi tamamlandıktan sonra, en yüksek puan alan sınırlayıcı kutular seçilmekte ve diğerleri elemine edilmektedir.

Literatürde yapılan çalışmalar [47], NMW yönteminin nesne tespiti ve nesne sınıflandırma alanında etkili olduğunu göstermektedir. Bu yöntem, çok sayıda nesnenin tespit edildiği durumlarda daha doğru sonuçlar elde etme potansiyeline sahiptir. Bu yöntem ayrıca, nesne takibi, nesne tanıma ve diğer benzeri uygulamalarda da kullanılabilmektedir.

1.3.3. Ağırlıklı Kutular Birleşimi

Ağırlıklı Kutular Birleşimi (Weighted Boxes Fusion, WBF) [13], nesne algılama sistemlerinin performansını artırmak üzere tasarlanmış bir sınırlayıcı kutu belirleme yöntemidir. Bu yöntem, birden fazla sayıdaki sınırlayıcı kutunun sonuçlarını daha doğru ve daha kararlı tek bir sınırlayıcı kutuda birleştirerek sonuç üretmeye odaklı çalışmaktadır. Bu çalışma; güven puanlarına dayalı olarak sınırlayıcı kutu koordinatlarının ilgili modelin ağırlığı baz alınarak birleştirilmesini içermektedir. Ağırlıklı kutuları birleştirme işlemi, birleştirilmiş kutuyu oluşturan sınırlayıcı kutu koordinatlarının ağırlıklı ortalamasının hesaplanması üzerine kurgulanmıştır. Bu ağırlıklar, ilgili sınırlayıcı kutuların güven puanları ile doğru orantılı olarak belirlenmektedir. Daha yüksek güven puanına sahip

sınırlayıcı kutuya daha yüksek ağırlıklar atayarak; güven puanı yüksek tahminlerden daha fazla etkilenen, düşük güven puanlı veya hatalı tahminlerden ise daha az etkilenen sonuç birleştirilmiş sınırlayıcı kutu üretilmektedir.

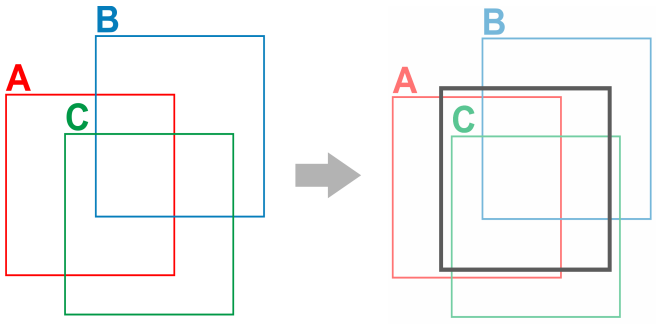
Ağırlıklı Kutular Birleşimi'nin amacı, farklı modellerin güçlü yanlarını birleştirerek daha güçlü bir sonuç elde etmektir. İlaveten, yanlış pozitiflerin veya aykırı değerlerin etkisini azaltarak nesne algılama modellerinin performansını artırmayı gaye edinmiştir. Bu tekniğin, Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS) gibi geleneksel yöntemlerle karşılaştırılması yapılmış ve işlem sonrası adımlarda daha etkili ve verimli olduğu geliştiricisi tarafından öne sürülmüştür.

1.3.4. Ortalama Alma

Ortalama Alma (OA), sınırlayıcı kutu koordinatlarının ortalamasının alınmasını benimseyen bir yöntemdir. İlgili yöntem; bütün sınırlayıcı kutulara eşit mesafede yaklaşmakta olup herhangi bir ağırlıklandırma durumu söz konusu değildir. Önerilen yöntem; girdi olarak verilen her bir sınırlayıcı kutunun sol üst/alt ($B_{sol_{alt/ust}}$) ve sağ üst/alt ($B_{sag_{alt/ust}}$) köşelerin piksel konumlarını çıkararak her birini kendi içerisinde ayrı ayrı incelemektedir. Şekil 1.8'de temsili çalışma şemasına yer verilen yöntemde, her bir sınırlayıcı kutu koordinatı için ilgili piksel değerinin toplam girdi olarak verilen sınırlayıcı kutu sayısına bölünmesi yolu ile ortalama alma işlemi uygulanmaktadır:

$$B_{sol_{alt/ust}}, B_{sag_{alt/ust}} = \frac{1}{n} \sum B_i, \tag{1.1}$$

İşlem sonrası üretilen piksel koordinatları birleştirilerek yeni bir sınırlayıcı kutu elde edilmektedir.



Şekil 1.8. Ortalama alma yöntemi çalışma şeması

Birbirine yakın konumlarda üretilmiş sınırlayıcı kutular üzerinde etkili sonuçlar veren yöntem; konumlandırmaları birbirinden uzak veya bağımsız sınırlayıcı kutular üzerinde etkili çalışamamaktadır. Bu sebeple; önerilen yöntemin kullanımı öncesinde; Jaccard (IoU) metriği ile hedef sınırlayıcı kutuların ön analizi sonrası, keşişim oranı 0.8 ve üzeri olan durumlarda kullanımı tavsiye edilmektedir.

2. BÖLÜM

YÖNTEM VE MATERYAL

Tez çalışmasının temel amacı, güncel yüz tespiti yöntemlerinin yakın kızılötesi görüntülerdeki başarımlarının araştırılması ve yeni sınırlayıcı kutu belirleme yöntemleri ile başarımlarının artırılmasıdır. Ancak literatürde, yakın kızılötesi görüntülerde yüz konumlarının sınırlayıcı kutularla etiketlendiği bir veri seti bulunmamaktadır. Bu sebeple; öncelikle TUFTS veri setindeki [2] yakın kızılötesi görüntüler etiketlenmiştir. Etiketleme ile elde edilen referans değerler (ground truth), hem yüz tespit yöntemlerinin hem de sınırlayıcı kutu belirleme yöntemlerinin başarımlarının araştırılmasında kullanılmıştır. Kullanılan veri seti, farklı kafa pozisyonlarında alınmış görüntüleri barındıran zengin bir içeriğe sahiptir. İlgili veri setine ilişkin detaylı bilgiler Bölüm 2.1.'de verilmiştir.

Gerçekleştirilen çalışmanın ana odağı olan yüz tespiti yöntemlerinin temeli nesne tespit yöntemlerine dayanmaktadır. Nesne tespit yöntemleri, son yirmi yılda devrim niteliğinde büyük bir ilerleme göstermiştir. Bu ilerlemenin derin öğrenme yöntemlerinin kullanılmasından kaynaklandığı bilinmektedir [20]. Bu sebeple derin sinir ağları tabanlı olarak geliştirilmiş ve başarımlarını kanıtlamış güncel yüz tespit yöntemleri ile çalışılmıştır. Bu yüz tespit yöntemleri sınırlayıcı kutu belirleme yöntemlerine girdi olarak verilerek daha başarılı yüz tespiti yapılması hedeflenmiştir. Hedef doğrultusunda kullanılan sınırlayıcı kutu belirleme yöntemlerinin detaylarına ve çalışma prensibine Bölüm 2.2'de yer verilmiştir. Ayrıca, Bölüm 1.3'te tanıtılan sınırlayıcı kutu birleştirme yöntemlerine ilaveten Bölüm 2.2'de iki yeni sınırlayıcı kutu belirleme yöntemi önerilmektedir.

2.1. TUFTS Veri Seti

TUFTS veri seti [2], çeşitli görüntü türleri ile kapsamlı ve büyük ölçekli bir biçimde yüz görüntülerini bünyesinde barındıran bir veri setidir. Bu kapsamlı veri seti; görünür görüntüleri, yakın kızılötesi görüntüleri, termal görüntüleri, bilgisayarlı eskizleri, LYTRO görüntüleri, kaydedilmiş videoları ve 3D görüntüleri içermektedir.

Veri seti, 15'ten fazla ülkeden 74 kadın ve 38 erkek ile 10.000'den fazla görüntüden oluşmaktadır. 4 ila 70 yaş aralığını kapsayan kişilerden toplanan veriler, ortalama 15 dakikalık çekimler sonucunda elde edilmiştir.

Zengin içeriği ve bünyesinde barındığı yakın kızılötesi görüntüler sebebiyle seçilmiş olunan TUFTS veri seti, farklı kafa pozisyonlarında alınmış görüntülere sahiptir. İlgili veri seti 2266 adet yan ve 1140 adet düz kafa pozisyonunda olmak üzere toplam 3406 adet yakın kızılötesi yüz görüntüsü içermektedir. Şekil 2.1, TUFTS veri setinde bulunan bazı yakın kızılötesi görüntü örneklerini göstermektedir. Bu veri setindeki tüm görüntüler bu çalışma kapsamında etiketlenerek, referans sınırlayıcı kutu konumları elde belirlenmiştir.



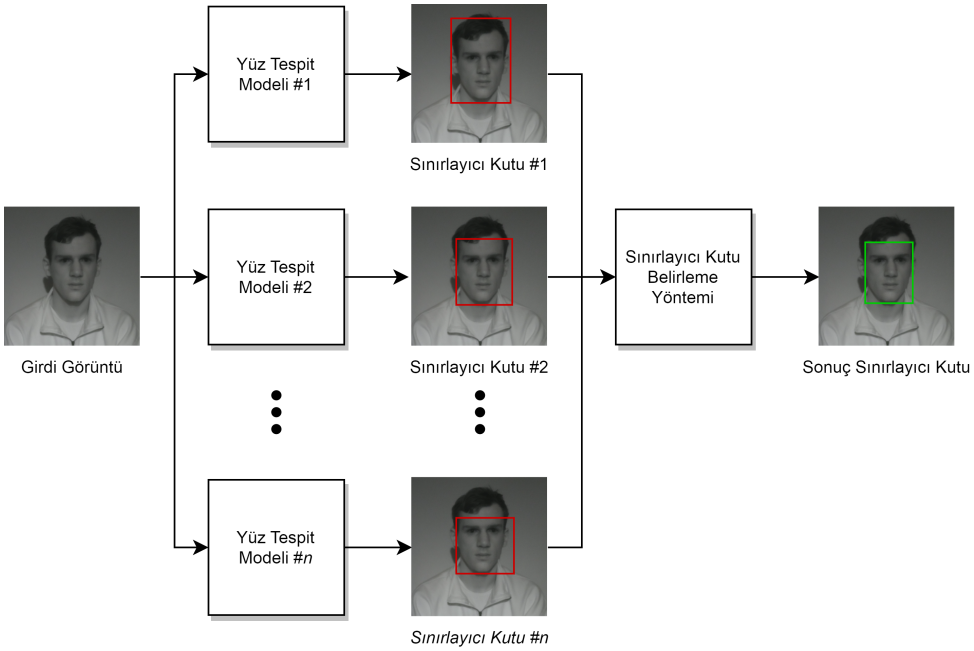
Şekil 2.1. TUFTS veri seti yakın kızılötesi örnek görüntüler

İlgili veri setinde yakın kızılötesi görüntüleme için aydınlatma, 850nm dalga boyuna sahip kızılötesi 96 LED ışık sistemi kullanılarak sağlanmıştır. Kameralar, katılımcının etrafında yaklaşık bir yarım daire oluşturacak şekilde dokuz eşit mesafeli konuma

getirilirken, her katılımcıdan sabit bir bakış açısına bakması istenerek görüntüler alınmıştır. Görüntü alınması esnasında her katılımcı, kameraya yakın mavi bir arka planın önünde oturtulmuştur. Kameralar, tripodlara monte edilerek her kameranın yüksekliği, görüntü merkezine karşılık gelecek şekilde manuel olarak ayarlanmıştır. Katılımcıya olan mesafe, görüntü alımı sürecinde sabit tutulmuş ve homojen ışık kullanılarak sabit bir aydınlatma koşulu sağlanması amaçlanmıştır.

2.2. Sınırlayıcı Kutu Belirleme Yöntemleri Kullanarak Yüz Tespiti

Sınırlayıcı kutu belirleme, birden fazla modelin tahminlerini beraber değerlendirerek daha başarılı sonuç sınırlayıcı kutu üretmek için kullanılan yöntemlerin genel adıdır [46]. Bu yöntemler, girdi olarak verilen her bir modelin daha başarılı olduğu yönünden yararlanmaktadır. Bu sayede, yapılan yüz tespitlerinin performansını artırıcı niteliktedir. Sınırlayıcı kutu belirleme yöntemlerinin çalışma yapısına Şekil 2.2'deki akış diyagramında yer verilmiştir. Şekilde görüleceği üzere; girdi görüntü her bir yüz tespit modeli tarafından analiz edilip ayrı ayrı yüz bölgeleri bulunmaktadır. Ardından, bütün yöntemlerin tespit etmiş oldukları sınırlayıcı kutu alanları, sınırlayıcı kutu belirleme yöntemine girdi olarak verilerek daha başarılı bir yüz konumu çıkarılmaktadır.



Şekil 2.2. Sınırlayıcı kutu belirleme yöntemi akış diyagramı

Sınırlayıcı kutu belirleme yöntemlerinin genel algoritması Algoritma 2.1'de yer almaktadır. Algoritmada görüleceği üzere, modeller tarafından tespit edilen yüz bölgelerinin sınırlayıcı kutu koordinatları girdi olarak alınmaktadır. Bu sınırlayıcı kutuların birbiriyle kesişim durumları karşılaştırılmaktadır. Eğer bir sınırlayıcı kutu ile diğerinin keşisimi varsa ilgili sınırlayıcı kutu belirleme yöntemi kullanılarak kutular beraber değerlendirilmektedir. Bu yolla yeni bir sınırlayıcı kutu üretilmektedir. Yapılan karşılaştırma işleminde, genellikle IoU metriği kullanılmaktadır. Belirlenen bir eşik değerin üzerindeki veya altındaki IoU değerine göre sınırlayıcı kutu belirleme yöntemleri çalışmaktadır. Sonuç olarak, sınırlayıcı kutu belirleme yöntemlerinin ürettiği yeni sınırlayıcı kutu koordinatları çıktı olarak verilmektedir.

Girdi: Sınırlayıcı kutuların listesi $B = \{B_1, B_2, \ldots, B_n\}$
Çıktı: Yeni üretilen sınırlayıcı kutular $B_{\text{sonuç}}$

Başlat: $B_{\text{sonuç}} = \emptyset$;
for $i = 1$ **to** n **do**
  if B_i *ile* $B_{sonuç}$ *içerisindeki bir sınırlayıcı kutunun* (B_j) *kesişimi* $\neq \emptyset$ **then**
    B_i ile B_j'yi beraber değerlendir;
  end
  else
    B_i'yi $B_{\text{sonuç}}$'a ekle;
  end
end
return $B_{\text{sonuç}}$;

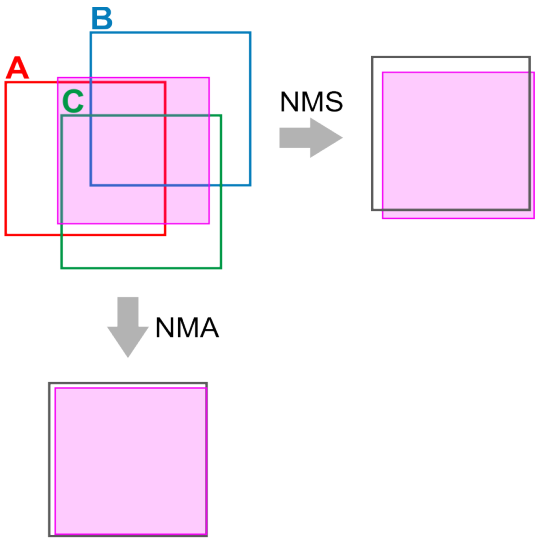
Algoritma 2.1: Sınırlayıcı kutu belirleme yöntemleri genel algoritması

Gerçekleştirilen tez çalışmasında; Bölüm 1.3'te verilen literatürdeki dört sınırlayıcı kutu belirleme yöntemi kullanılarak, girdi sınırlayıcı kutuların iyileştirilmesi ve bu yöntemlerin performanslarının analizi amaçlanmıştır. Buna ilaveten; Maksimum Olmayan Ortalama ve Kombinasyonel Oylama olarak isimlendirilmiş iki yeni sınırlayıcı kutu belirleme yöntemi bu bölümde önerilmektedir. Maksimum Olmayan Ortalama, en yüksek güven skoruna sahip sınırlayıcı kutu ile diğer sınırlayıcı kutuların her birini analiz etmekte ve ortalama alma yaklaşımı ile yeni bir sınırlayıcı kutu üretmektedir. Kombinasyonel Oylama ise, sınırlayıcı kutu topluluğu kombinasyonlarının oylama yöntemi temel alınarak değerlendirilmesine dayanmaktadır. Bu yöntem genel manada, farklı modellerinin tahminlerini bir araya getirerek en çok tercih edilen konumda yeni bir sınırlayıcı kutu oluşturmaktadır. Önerilen yöntemlerin detayları alt başlıklarda yer almaktadır.

2.2.1. Maksimum Olmayan Ortalama

Önerilen ilk sınırlayıcı kutu belirleme yöntemi Maksimum Olmayan Ortalama (Non-Maximum Averaging, NMA) yöntemidir. Bu yöntem, Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS) ve Maksimum Olmayan Ağırlıklandırma (Non-Maximum Weighting, NMW) yöntemlerinden esinlenilerek geliştirilmiştir.

NMS, en yüksek güven skorlarına sahip sınırlayıcı kutuyu seçmek üzere kurgulanmış bir yöntemdir. Bu yöntem, en yüksek güven skorlarına sahip olan sınırlayıcı kutu ile diğer her bir sınırlayıcı kutunun IoU metriğine göre karşılaştırmasını yapmaktadır. Elde edilen sonuç değer, eşik değerden daha yüksek ise ilgili sınırlayıcı kutu bastırılmaktadır. Bu bastırma işlemleri neticesinde, en yüksek güven skoruna sahip sınırlayıcı kutu sonuç olarak çıkarılmaktadır. NMW ise; NMS yönteminin gerçekleştirdiği bastırma işlemini ağırlıklandırma faktörlerini kullanarak yapmaktadır. Bu faktörler; hangi sınırlayıcı kutuların saklanacağını veya bastırılacağını belirlemede aktif rol oynamaktadır. Bu sayede, NMS yönteminin güven skorları üzerine türetmiş olduğu yaklaşıma ağırlık faktörü eklenerek yeni bir yöntem öne sürülmüştür. Buradaki yaklaşımdan esinlenilerek NMA yöntemi önerilmektedir. Önerilen yöntem, NMS yönteminin en yüksek güven skorlu sınırlayıcı kutuyu seçme stratejisini ortalama alma tekniğiyle ileriye taşımaktadır. İlaveten, NMW yönteminin ağırlık belirleme sırasındaki işlem süresinin de önüne geçerek daha basit ve etkili bir yaklaşım sunmaktadır. Sunulan yöntemin NMS yöntemi ile karşılaştırmalı tespit bölgeleri Şekil 2.3'teki şemada yer almaktadır.



Şekil 2.3. NMS ve NMA yöntemleri karşılaştırma şeması

Önerilen yöntem kapsamında, öncelikle NMS yönteminde olduğu gibi en yüksek güven skoruna sahip sınırlayıcı kutu (B_{maks}) belirlenmektedir. Belirlenen bu sınırlayıcı kutu, diğer tüm sınırlayıcı kutular ile IoU performans değerlendirme metriği kullanılarak karşılaştırılmaktadır:

$$I_i = IoU(B_i, B_{maks}) \tag{2.1}$$

burada (I_i), i. sınırlayıcı kutu (B_i) ile en yüksek güven skoruna sahip sınırlayıcı kutunun (B_{maks}) IoU metrik değerini ifade etmektedir.

Ardından; karşılaştırma neticesinde elde edilen değer kullanılarak, ilgili sınırlayıcı kutunun koordinat değerleri üzerinde ortalama alma işlemi uygulanmaktadır. Sonuç sınırlayıcı kutu (B_{sonuc}), takip eden denklemde verilen ortalama yaklaşımı ile elde edilmektedir:

$$B_{sonuc} = \frac{\sum_{i=1}^{n} B_i \times I_i}{\sum_{i=1}^{n} I_i} \tag{2.2}$$

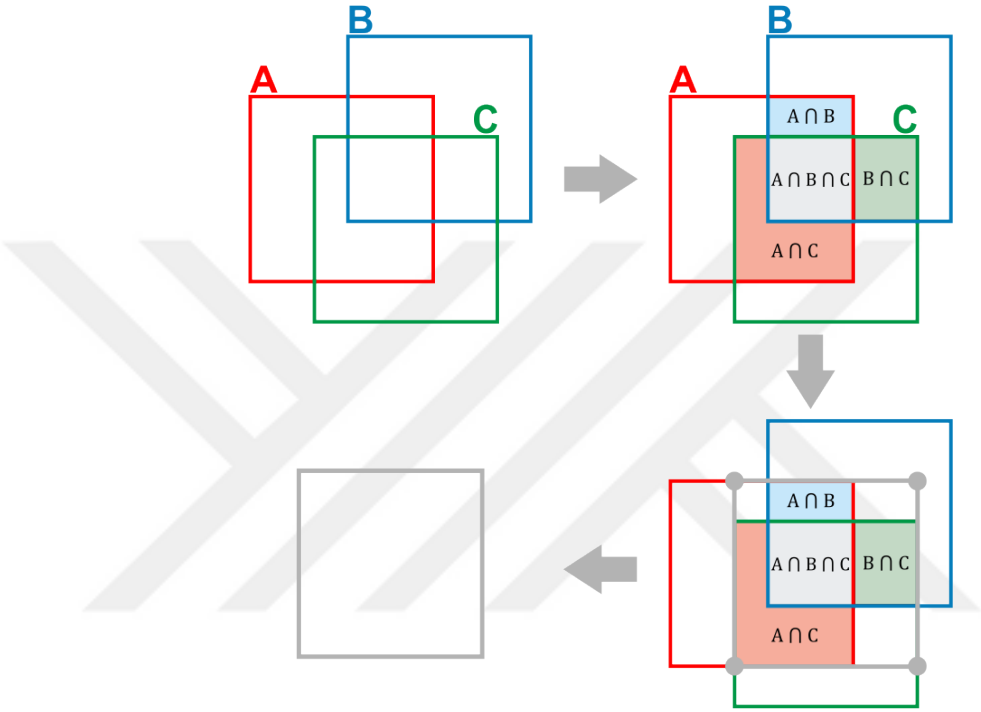
2.2.2. Kombinasyonel Oylama

Tez çalışması kapsamında önerilen ikinci sınırlayıcı kutu berlirleme yöntemi Kombinasyonel Oylama yöntemidir. Kombinasyonel Oylama, sınırlayıcı kutu topluluğu kombinasyonlarının oylama yöntemi temel alınarak değerlendirilmesine dayanmaktadır. Bu yöntem, farklı modellerinin tahminlerini bir araya getirerek en çok tercih edilen tahmini seçmek üzere tasarlanmıştır.

Oylama işleminde karar yetkisi çoğunluğa aittir. Burada bahsedilen çoğunluk, mevcut katılımcı sayısının yarısından fazla oya sahip tarafı ifade etmektedir. Önerilen sınırlayıcı kutu belirleme yönteminde, bu çoğunluğun belirlenmesi sürecinde olası kombinasyonların analizine dayalı bir yaklaşım sergilenmektedir.

Örneğin; Şekil 2.4'te görüldüğü gibi üç adet sınırlayıcı kutunun mevcut olduğu bir problemde, en az iki sınırlayıcı kutunun aynı alanda yüz bölgesi olduğunu tespit etmesi gerekmektedir. Yani, yüz tespiti için ya kutulardan iki tanesinin ya da üç tanesinin seçim alanlarını barındıran bölgeler aranmaktadır. Bu arama sürecinde, Şekil 2.4'te görüleceği üzere A sınırlayıcı kutusu, B ile olduğu gibi C ile de ortak noktalara sahiptir. Bu durum,

farklı kombinasyonlara bakılması gerektiğini göstermektedir. Bu sebeple; girdi olarak verilen sınırlayıcı kutuların sayısından başlanarak, sınırlayıcı kutu sayısının yarısına kadar olası bütün kombinasyonlar hesaplanmaktadır. Hesaplama neticesinde, elde edilen kombinasyonların içeriğindeki sınırlayıcı kutuların ortak olarak karşılık geldiği bölgeler kaydedilerek her birinin birleşimi alınıp sonuç sınırlayıcı kutu elde edilmektedir.



Şekil 2.4. Kombinasyonel oylama yöntemi çalışma şeması

Önerilen yöntemde, kombinasyona dayalı oylama işlemi yapılmadan önce iki defa eleme işlemi uygulanmaktadır. Bu işlemlerden ilki alakasız sınırlayıcı kutuları elimine etmek üzeredir. İlgili eleme işleminde, girdi olarak verilen her bir modele ait sınırlayıcı kutular birbiri ile ayrı ayrı IoU metriği baz alınarak karşılaştırılmaktadır:

$$IoU(B_i, B_j) = \frac{B_i \cap B_j}{B_i \cup B_j} \tag{2.3}$$

Karşılaştırma neticesinde elde edilen sonuç değer, belirlenen eşik değerden büyük ise bir listeye eklenmektedir. Bu çalışmada eşik değer 0,7 olarak seçilmiştir. Listeye eklenmiş olan sınırlayıcı kutular benzeri konumlarda yer alan sınırlayıcı kutular olup aykırı konumlardaki sınırlayıcı kutular elenmektedir.

Gerçekleştirilen ikinci eleme işlemi, oylama için karar verici sınırlayıcı kutuları belirlemek için yapılmaktadır. Yapılan bu işlemde, önceki eleme işleminde oluşturulan liste içerisindeki sınırlayıcı kutular kullanılmaktadır. Bu süreçte, öncelikle liste içerisindeki en büyük alana sahip sınırlayıcı kutu (B_{maks}) belirlenmektedir. Belirlenen bu sınırlayıcı kutu, takip eden denklemde görüleceği gibi liste içerisindeki diğer sınırlayıcı kutular ile IoU metriği ile analiz edilmektedir:

$$IoU(B_i, B_{maks}) = \frac{B_i \cap B_{maks}}{B_i \cup B_{maks}} \tag{2.4}$$

Analiz neticesinde bulunan değer, belirlenen eşik değerden küçük ise listeden çıkarılmaktadır. Bu çalışmada eşik değer 0,8 olarak belirlenmiştir. Bu sayede, farklı kombinasyonlar ile daha yüksek doğruluk oranına sahip sonuç sınırlayıcı kutunun üretildiği oylama yöntemine katkı sağlayan karar verici sınırlayıcı kutular belirlenmektedir.

Belirleme aşamasında gerçekleştirilen ilk eleme işlemi ile hedef nesne ile alakasız sınırlayıcı kutular göz ardı edilmektedir. Temel odak alanının net olarak belirlendiği bu süreçten sonra; o alana en çok hâkim olan, en geniş yer kaplayan sınırlayıcı kutu seçilerek onun etrafında diğer sınırlayıcı kutuların analizi ile ikinci bir eleme işlemi daha gerçekleştirilmektedir. Diğer başlıklarda verilen yöntemlerde, analizlerin gerçekleştirilmesi için en yüksek güven skorlarına sahip sınırlayıcı kutu seçilmektedir. Ancak, farklı modeller farklı güven skorları üretmektedir. Şekil 3.4'te görüleceği üzere, modeller kendi içlerinde çok farklı güven skorlarına sahiptirler. Diğer taraftan, tespit noktası daha doğru olan bir modelin kendi güven skorunun daha düşük olduğu da görülebilmektedir. Bu sebeple, bu yanıltıcı duruma çözüm önerisi getirmek için güven skorlarından bağımsız ve tarafsız bir yaklaşım önerilmektedir. İlgili yaklaşım, gelişmiş konumlandırma ve yüksek doğruluk oranlı yeni sınırlayıcı kutu üretme konusunda umut vericidir.

3. BÖLÜM

BULGULAR

Birçok uygulama alanında yüz tespitinin aydınlatma durumundan bağımsız olarak verimli bir şekilde gerçekleştirilebilmesi gerekmektedir. Bu sebeple, görünür ışığın yeterli olmadığı durumlarda dahi görüntülenen sahne hakkında bilgi sağlayabilen yakın kızılötesi kameraların kullanımı önem kazanmaktadır. Ancak, literatürde yakın kızılötesi görüntüler üzerinde yüz tespitine yönelik oldukça kısıtlı sayıda çalışma bulunmaktadır. Hem bu tespit edilen açığı kapatmak hem de yüz tespiti verimliliğini artırabilmek amacıyla derin sinir ağı tabanlı yüz tespiti yöntemlerinin yakın kızılötesi görüntüler üzerinde karşılaştırmalı analizi yapılmıştır. Bölüm 3.3'te verilen ilgili karşılaştırma çalışması kapsamında, görsel ve nicel olmak üzere karşılaştıma neticeleri sunulmuştur.

Kullanılan modellerin performanslarının değerlendirilebileceği bir diğer karşılaştırma yaklaşımı işlem zamanı karşılaştırmasıdır. Bu yaklaşımda, aynı yakın kızılötesi görüntüler üzerinde çalıştırılan modellerin aynı işlemi gerçekleştirmek için harcamış oldukların sürelerin analizi yapılmıştır. Bu çalışma hem CPU hem de GPU üzerinde yürütülerek ayrı ayrı analiz sonuçları elde edilmiştir. Buna ilaveten, kullanılan modellerin sahip oldukları parametre sayıları, FLOP miktarları ve omurga ağları değerlendirilmiş olup bütün içerikler Bölüm 3.3.3'te karşılaştırmalı olarak verilmiştir.

Sınırlayıcı kutu belirleme yöntemleri, birden fazla sınırlayıcı kutunun beraber analiz edilmesi ile daha isabetli tespit noktalarının üretilmesini amaçlayan yaklaşımlara sahiptirler. Buna dayanarak, yakın kızılötesi görüntülerde başarımları analiz edilen yüz tespiti modellerinin, sınırlayıcı kutu belirleme yöntemlerinin kullanımı ile performanslarının iyileştirilmesi üzerine çalışılmıştır. Bu bağlamda, literatürde mevcut olan dört farklı sınırlayıcı kutu belirleme yöntemi ve ilgili tez çalışması dahilinde sunulan

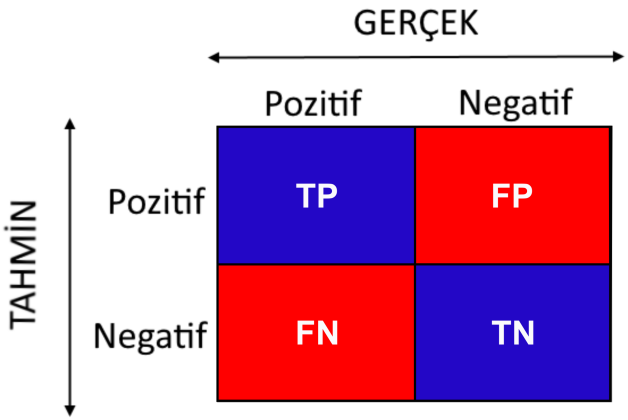
iki farklı yeni yöntem kullanılmıştır. Bu yöntemlerin karşılaştırmalı analizlerine ilişkin görsel ve nicel sonuçlar Bölüm 3.4'de yer almaktadır.

3.1. Deney Altyapısı

Bu tez çalışması kapsamında yapılan tüm deneyler, Intel Core i9 10850K 3.60 GHz 20MB önbelleğe sahip işlemciye, RTX 3060 12GB 192 bit grafik işlemciye ve 64 GB RAM belleğe sahip bilgisayar üzerinden gerçekleştirilmiştir.

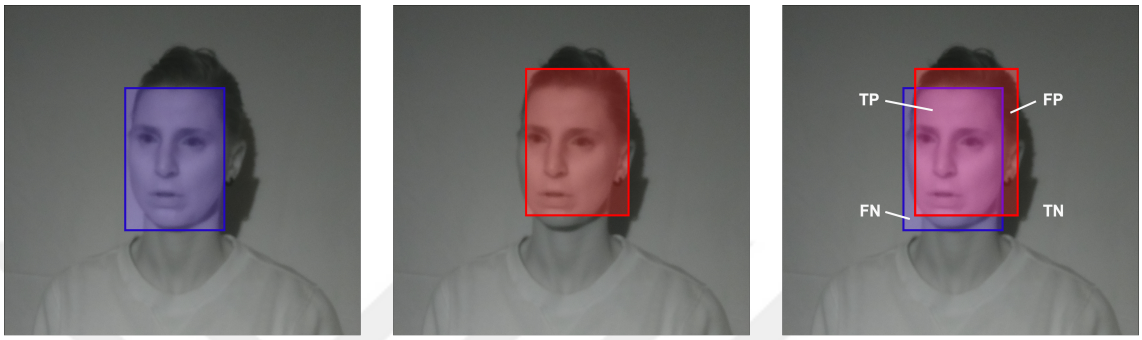
3.2. Performans Değerlendirme Metrikleri

Performans değerlendirme metrikleri, bir modelin veya sistemin performansını ölçmek ve değerlendirmek için kullanılan ölçütlerdir. Bu metrikler, farklı problemler ve gereksinimler için farklı sonuçlar üretebilmektedirler. Bu nedenle farklı performans değerlendirme metrikleri kullanılmıştır. Kullanılan metrikler, gerçek değerler ile modeller tarafından tahmin edilen değerlerin karşılaştırılması suretiyle çalışmaktadır. Bu çerçevede, Şekil 3.1'de verilen karmaşıklık matirisi gerçek değerlerle tahmin edilen değerler arasındaki ilişkiyi göstermektedir. TP (True Positive); gerçek pozitif değerlerin doğru bir şekilde tahmin edildiğini, FP (False Positive) ise gerçek negatif değerlerin yanlış tahmin edildiğini ifade etmektedir. Diğer taraftan, TN (True Negative); gerçek negatif değerlerin doğru bir şekilde negatif olarak tahmin edildiği, FN (False Negative) ise gerçek pozitif değerlerin yanlış tahmin edilmesi durumunu ifade etmektedir.



Şekil 3.1. Karmaşıklık matrisi

Şekil 3.2 incelendiğinde, karmaşıklık matrisinde verilen değerlerin örnek bir yüz tespiti işlemindeki karşılıkları görülebilmektedir. Şekil 3.2 (a), referans değer ile oluşturulmuş gerçek yüz bölgesini, Şekil 3.2 (b), yüz tespit yöntemi tarafından tahmin edilen yüz bölgesini, Şekil 3.2 (c) ise (a) ve (b)'nin beraber değerlendirildiği kesişim alanları doğrultusunda elde edilen TP, FN, FP ve TN alanları ifade etmektedir.



Şekil 3.2. Örnek yüz tespiti: (a) gerçek yüz bölgesi, (b) tahmin edilen yüz bölgesi, (c) yapılan tahmin için TP, FN, FP ve TN alanlar

Gerçekleştirilen çalışmaların değerlendirilmesi sürecinde, beş farklı performans değerlendirme metriği seçilmiş ve yapılan bütün çalışmalarda bu metrikler kullanılmıştır. Seçilen performans değerlendirme metrikleri; Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Skor (F1 Score) ve Jaccard (IoU) olup her biri aşağıda açıklanmıştır:

Doğruluk

Doğruluk metriği, bir modelin performansını değerlendirmek için kullanılan popüler ölçütlerdendir. Bu ölçüt, toplam veri sayısının ne kadar doğru tespit edilebildiğini bulmaya odaklanmaktadır. Bunun için, doğru yapılan tahmin sayısının toplam tahmin sayısına oranlanması ile hesaplanmaktadır:

$$Doğruluk = \frac{TP + TN}{TP + FP + TN + FN}. \tag{3.1}$$

Doğruluk, özellikle dengesiz sınıf dağılımlarında (imbalanced class distribution) yanıltıcı olabilmektedir. Bu sebeple, diğer başlıklarda verilen metrikler ile beraber kullanılmıştır.

Kesinlik

Genellikle sınıflandırma problemlerinde kullanılan bir performans değerlendirme metriğidir. Kesinlik, en basit ifadeyle gerçek pozitifler (TP) ile tüm pozitiflerin (TP, FP) oranıdır:

$$Kesinlik = \frac{TP}{TP + FP}. \tag{3.2}$$

Duyarlılık

Genellikle sınıflandırma problemlerinde kullanılan bir diğer metrik Duyarlılıktır. Gerçek pozitiflerin (TP) tüm doğru tespitlere (TP, FN) oranını göstermektedir:

$$Duyarlılık = \frac{TP}{TP + FN}. \tag{3.3}$$

F1 Skor

F1 Skor; Kesinlik ve Duyarlılık performans değerlendirme metriklerinin harmonik ortalamasının hesaplanması ile elde edilmektedir:

$$F1Skor = \frac{2TP}{2TP + FP + FN} = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık}. \tag{3.4}$$

Hem Kesinlik hem de Duyarlılığın ürettiği değeri dikkate alarak dengeyi sağlamaktadır.

Jaccard

Jaccard (Intersection over Union, IoU), iki kümenin benzerliğini ölçmek için kullanılan bir performans değerlendirme metriğidir. Bu metrik, iki kümenin kesişim kümesinin, birleşim kümesine oranını hesaplayarak benzerlik derecesini belirlemeye odaklı olarak çalışmaktadır. Jaccard benzerlik katsayısından ismini alan bu metrik, 0 ile 1 arasında bir değer almaktadır. 0, iki kümenin hiç ortak noktasının olmadığını, 1 ise iki kümenin tamamen kesişmekte olduğunu göstermektedir.

$$Jaccard(Nesne_1, Nesne_2) = \frac{Nesne_1 \cap Nesne_2}{Nesne_1 \cup Nesne_2}, \tag{3.5}$$

burada; $Nesne_1 \cap Nesne_2$ verilen nesnelerin keşisim miktarlarını, $Nesne_1 \cup Nesne_2$ ise birleşim miktarlarını temsil etmektedir. Oranları ise Jaccard benzerlik katsayısını oluşturmaktadır.

3.3. Yüz Tespiti Derin Sinir Ağı Modellerinin Karşılaştırılması

Derin sinir ağı modelleri, büyük miktarda veriyi kullanarak karmaşık özellikleri öğrenme kabiliyetine sahiptir. Bu kabiliyetlerinden dolayı, görüntülerdeki yüzleri tespit etmek ve tanımak için yaygın olarak kullanılırlar. Derin sinir ağı modelleri ile geliştirilmiş yüz tespit yöntemleri arasında performans ve hassasiyet farklılıkları olmaktadır. Belli bir senaryo üzerinde bir model iyi çalışırken, başka bir durumda diğer bir model daha yüksek başarı oranı ile yüz tespiti yapabilmektedir. Bu sebeple, bu bölümde, literatürde eksikliği tespit edilmiş kızılötesi görüntüler üzerinde yüz tespiti senaryosunda, farklı yüz tespit yöntemlerinin karşılaştırmalı analizi yapılmıştır. Literatürde yaygın olarak kullanılan popüler sekiz yüz tespiti yöntemi kızılötesi yüz görüntüleri içeren TUFTS veri seti içerisindeki, düz ve yan kafa pozisyonlarındaki görüntüler üzerinde çalıştırılmıştır. Elde edilen sonuçlar; Doğruluk, Kesinlik, Duyarlılık, F1 Skor ve Jaccard performans değerlendirme metrikleri kullanılarak değerlendirilmiştir.

3.3.1. Yüz Tespiti Modellerinin Görsel Karşılaştırması

Yüz tespiti derin sinir agı modellerinin yakın kızılötesi görüntüler üzerinde görsel karşılaştırılması yoluyla, elde edilen sonuçlar somutlaştırılmıştır. Verilen görsellerin yorumlaması doğrultusunda çıkarılan analizler, ilgili yöntemlerin karşılaştırmalı performanslarını gösterir niteliktedir.

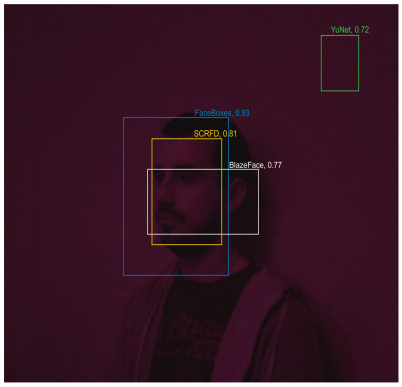
Bazı yüz tespit yöntemleri tarafından tespit edilen yüz bölgelerinin karşılaştırmalı ele alındığı görsellere Şekil 3.3'te yer verilmiştir. Görsellerde görüleceği üzere, bazı modeller yüz alanının dışındaki farklı konumları da yüz bölgesi olarak tanıyabilmekte veya mevcut yüz alanının belli kısımlarını tespit edememektedir. Yöntemlerin başarı oranlarını direkt etkileyen bu durum, yüz tespiti sonrası gerçekleştirilebilecek yüz tanıma, kafa pozu tahmini, duygu analizi, yüz bölgesindeki göz, burun, ağız vb. uzuvların tespiti gibi problemlere getirilen çözüm yaklaşımlarını da doğrudan etkilemektedir. Buna dayanarak, tişörtteki yüz figürünün veya duvara yansıyan gölgenin yüz bölgesi olarak algılanması, sonraki aşamalarda üzerine çalışılacak bu gibi tespitleri zora sokacak niteliktedir. Bu sebeple, yüz tespit yöntemlerinin karşılaştırmalı analizi kafa pozisyonlarına da bakılarak

ayrı ayrı yapılmıştır. Değerlendirme sonuçları ve örnek görseller beraber incelendiğinde, bulanık bir görüntüde dahi yüz bölgesini doğru tespit eden SCRFD modelinin en yüksek doğruluk oranlarına, YuNet modelinin ise en düşük doğruluk oranlarına sahip olması durumu anlaşılmaktadır.



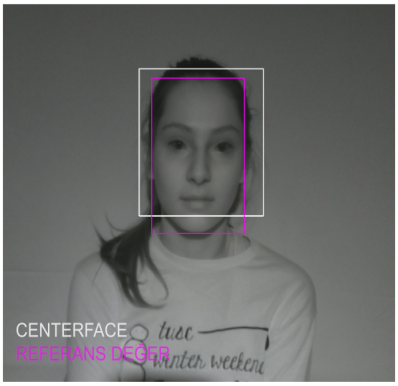
Şekil 3.3. Farklı yüz tespit yöntemleri tarafından tespit edilen yüz bölgeleri

Literatürde yer alan her yüz tespit yönteminin başarı değerlendirme kriterleri ve kullandığı referans değerler birbirinden farklıdır. Bu sebeple, kendi algoritmaları tarafından tespit doğruluk oranlarını tahmin etmek üzere ürettikleri güven skorları farklılık göstermektedir. Bu durum, yüz tespit yöntemlerinin ürettiği güven skorlarının yer aldığı Şekil 3.4'te daha iyi görülebilmektedir. Burada; %93 ile en yüksek güven skoru üreten FaceBoxes, yüz bölgesi dışında boyun bölgesini ve diğer alakasız bölgeleri de yüz alanı olarak tespit ederken, doğru yüz bölgesine en yakın konumlandırmayı yapan SCRFD, %81 oranında güven skoru üretmiştir. Diğer taraftan; yanlış bir yüz bölgesi tespiti yapan BlazeFace %77, yüz bölgesi ile alakası olmayan bir alanı seçen YuNet ise %72 oranında güven skoruna sahiptir. Dolayısıyla, yüz tespit yöntemlerinin üretmiş oldukları güven skorları tutarlılık göstermemektedir. Bu sebeple, gerçekleştirilen çalışmalar kapsamında elde edilen doğruluk oranları, etiketleme ile oluşturulmuş referans değerler temel alınarak çıkarılmış ve bu sıkıntılı durumun önüne geçilmiştir.



Şekil 3.4. Yüz tespit yöntemleri tarafından bulunan yüz bölgeleri ve ürettikleri güven skorları

Yüz tespiti sonrası örnek bir sonuç görüntüsüne Şekil 3.5'te yer verilmiştir. Burada, etiketleme işlemi sonrasında üretilen referans değer ve CenterFace modeli tarafından tespit edilen yüz bölgeleri görülmektedir. İlgili şekilde görüleceği üzere, referans değer tam yüz bölgesini çerçevelemekteyken, yüz tespit yönteminin tespit noktası çene kısmını alamamakta ve yüz alanını enlemesine daha geniş olarak bulmaktadır. Bu örnekten yola çıkılarak, çene bölgesine odaklı gerçekleştirilecek bir çalışmada, kullanılan yüz tespit yöntemi başarısız olacaktır. İlaveten, tespit edilen daha geniş yüz bölgesi, bu alanda gerçekleştirilecek uzuv (göz, ağız, vs.) tespiti gibi başka işlemlerin maliyetini artıracak niteliktedir. Bu sebeple, yüz bölgesinin net olarak tayin edilmesi büyük önem taşımaktadır. Bu doğrultuda, renkli ve gri seviye görüntülerde başarılı sonuçlar veren yüz tespit yöntemlerinin yakın kızılötesi görüntüler üzerindeki performansları analiz edilerek, yetersiz aydınlatma durumlarında yüz tespiti işleminin yapılmasının elzem olduğu durumlarda işlem maliyetini düşürecek ve yüz tespiti sonrasında gerçekleştirilecek işlemlerin başarısını artıracak bir yaklaşım sergilenmiştir.



Şekil 3.5. Referans değer ve CenterFace tarafından çıkarılan yüz bölgeleri

Tez çalışması kapsamında kullanılan bütün yüz tespit yöntemleri, belirlenen bir yakın kızılötesi görüntü üzerinde ayrı ayrı çalıştırılarak tespit ettikleri yüz bölgeleri analiz edilmiştir. Yapılan analizlere ilişkin görsel, Şekil 3.6'da verilmiştir. Verilen şekilde görüleceği üzere, referans değer ve dokuz farklı yüz tespit yönteminin üretmiş oldukları yüz bölgesi alanları ayrı ayrı mevcuttur.

Şekil 3.6 (a), etiketleme işlemi neticesinde elde edilmiş yüz bölgesini göstermektedir. Bu alan; alın, kulak, çene sınırları ve kısmen yan bir poz olması sebebiyle kaş,yanak sınırlarını kapsayacak şekilde işaretlenmiştir. İlgili görselde görüleceği üzere, yüz bölgesini net bir şekilde almaktadır.

Şekil 3.6 (b), SCRFD modelinin tespit ettiği yüz bölgesi alanını göstermektedir. SCRFD, kulak bölgesi ile yanak arasındaki alanı kısmen alamamış ve alın bölgesinden ufak taşma yaparak baş kısmını da işaretlemiştir. Ancak genel olarak başarılı bir yüz bölgesi tespiti yapmış olduğu görülmektedir.

Şekil 3.6 (c), RetinaFace modeli tarafından tespit edilen yüz bölgesi alanına aittir. Bu görselde görüleceği üzere RetinaFace, kişinin kulağı ile yanağı arasında kalan bölgeyi alamamıştır. Ayrıca, alın bölgesinden ve kaş hizasından kısmi taşma olacak şekilde işaretleme yaptığı görülmektedir. Genel yüz bölgesi alanına bakıldığında, RetinaFace modelinin başarılı bir yüz tespiti yaptığı söylenebilmektedir.

Şekil 3.6 (d), DSFD modeli tarafından tespit edilen yüz bölgesini göstermektedir. DSFD modeli, kulak ve yanak arasında kalan bölgeyi tespit edememiştir. İlaveten, baş bölgesine taşacak şekilde bir yüz alanı oluşturduğu görülmektedir. Genel tespit alanına bakıldığında, başarılı bir şekilde kişinin yüz bölgesini bulduğu görülmektedir.

Şekil 3.6 (e)'de, CenterFace modelinin tespit etmiş olduğu yüz bölgesi görülmektedir. Şekilde görüleceği üzere, yüz alanı oldukça geniş bir şekilde bulunmuştur. Çene, kulak, alın ve yanak bölgelerinden taşmalar mevcut olup boğaz, saç ve arka fon alanları dahi yüz bölgesi alanına dahil edilmiştir. CenterFace tarafından yapılan tespit, yüz alanının tamamını alıyor olsa dahi yüz tespiti işlemi sonrasında bu alan üzerinde çalışılabilecek olası diğer işlemlerin maliyetini artırabilecek niteliktedir.

Şekil 3.6 (f), FaceBoxes modeli tarafından tespit edilen yüz bölgesini göstermektedir. Tespit edilen alan; kulak, çene, yanak ve alın bölgelerinden taşmalara sahiptir. Kullanılan model, yüz bölgesi alanının tamamını işaretlemiştir. Ancak yüz alanının dışındaki kısımları da bu alana dahil ettiği görülmektedir. Bu sebeple, tespit edilen alan içerisinde yapılması muhtemel işlemlerin maliyetini artırabilecek yapıdadır.

Şekil 3.6 (g), RetinaFace'in MobileNet omurga ağı kullanılarak eğitilmiş modeli tarafından tespit edilmiş yüz bölgesi alanını göstermektedir. İlgili şekilde, kulak ile yanak arası bölgenin tespit edilemediği görülmektedir. Ayrıca, kaş, yanak ve alın bölgelerinden taşmalara sahip bir şekilde yüz bölgesi tespiti yapılmıştır.

Şekil 3.6 (h), BlazeFace modeli tarafından tespit edilen yüz bölgesi alanına aittir. Görselde görüleceği üzere; kullanılan model alın bölgesini tam alamamıştır. Yanak, kaş, çene ve kulak bölgelerinden oldukça fazla taşmalar bulunmaktadır. İlaveten; boğaz bölgesi, arka fon alanı ve kişinin giymiş olduğu tişörtün boyun alanı dahi yüz bölgesi olarak işaretlenmiştir. Genel olarak bakıldığında, yapılan tespit işleminin başarılı bir işlem olduğu söylenememektedir.

Şekil 3.6 (k), Dlib tarafından tespit edilen yüz bölgelerini göstermektedir. Verilen görselde görüleceği üzere; üç farklı yüz bölgesi tespit edilmişir. Bunlardan ilki; gerçek yüz bölgesine en yakın olan yüz tespiti alanıdır. Bu alanda; kişinin alın bölgesinin ve yanak ile kulak bölgesi arasındaki kısmın tam alınamadığı görülmektedir. Diğer taraftan, arka fon alanı ve saç bölgeleri yüz alanına dahil edilmiş olup işaretlenen alanda taşmalar mevcuttur. Tespit edilen ikinci yüz alanı; kişinin kabarık saçı ve saçının oluşturmuş olduğu gölgeyi kapsayacak şekilde işaretlenmiştir. Tespit edilen üçüncü yüz bölgesi ise; kişinin kolu ile tişörtü arasında kalan gölgeyi göstermektedir. Tespit edilen ikinci ve üçüncü yüz bölgelerinin tamamen yanlış olduğu ve birinci tespit bölgesinin de hatalı olduğu ilgili şekilde görülmektedir. Bu bağlamda, Dlib'in yakın kızılötesi görüntülerde başarısız bir yöntem olduğu sonucu çıkarılabilmektedir.

Şekil 3.6 (l), YuNet modeli tarafından tespit edilmiş yüz bölgelerine aittir. Yunet, verilen görüntüde beş farklı yüz bölgesi tespit etmiştir. Bunlardan birincisi, gerçek yüz alanı ile kısmen örtüşmektedir. Burada, yanak bölgesi ve alın bölgesinin çoğu tespit edilememiştir. Çene bölgesinin ise kısmen alınamadığı görülmektedir. Tespit edilen ikinci yüz bölgesi,

kişinin saçı ve arkasına düşen gölgesidir. Üçüncü yüz bölgesi, yine kişinin kabarık saçıdır. Dördüncü yüz bölgesi, kişinin giymekte olduğu tişörtün üzerindeki figürlerdir. Beşinci yüz bölgesi ise kişinin kolu ile tişörtü arasında oluşmuş gölgeye karşılık gelmektedir. Sonuç olarak; birinci tespit alanının hatalı, diğer dört tespitin ise yanlış olduğu verilen şekilde açık olarak görülmektedir. Bu bağlamda YuNet modelinin, yakın kızılötesi görüntülerde başarısız bir performansa sahip olduğu sonucu çıkarılabilmektedir.



Şekil 3.6. Kullanılan bütün yüz tespit yöntemleri tarafından aynı görüntüde yüz tespiti: (a) referans yüz bölgesi, (b) SCRFD, (c) RetinaFace, (d) DSFD, (e) CenterFace, (f) FaceBoxes, (g) RetinaFace(MobileNet), (h) BlazeFace, (k) Dlib, (l) YuNet

3.3.2. Yüz Tespiti Modellerinin Sayısal Karşılaştırması

Yüz tespiti işlemi, aydınlatma durumundan bağımsız bir şekilde gerçekleştirilebildiği gibi farklı kafa pozisyonlarında da verimli çalışabilmelidir. Bu nedenle, yakın kızılötesi görüntülerden oluşturulmuş TUFTS yüz görüntüsü veri setindeki görüntüler ayrıştırılarak düz ve yan kafa pozlarının ayrı ayrı değerlendirilmesi hedeflenmiştir. Düz kafa pozisyonundaki resimlerden elde edilen sonuç değerleri Tablo 3.1'de, yan kafa pozisyonundaki sonuç değerleri Tablo 3.2'de ve tamamının beraber ele alındığı değerler ise Tablo 3.3'te yer almaktadır. Tablolar incelendiğinde, yan kafa pozisyonlarında gerçekleştirilen yüz tespiti işlemlerindeki başarının düz kafa pozisyonuna göre daha düşük olduğu görülmektedir. Bu durum, yüz tespit yöntemlerinin bütün görüntüler üzerindeki başarısını da doğrudan etkileyerek farklı kafa pozisyonlarına karşı tutarlılıklarını ortaya koymaktadır. Örneğin; MobileNet kullanılarak eğitimi gerçekleştirilmiş RetinaFace yüz tespit yöntemi; düz kafa pozisyonlarında %89 doğruluk oranına sahipken, yan kafa pozisyonlarında bu oran %64'e düşmektedir. Aradaki bu büyük fark, ilgili yüz tespit yönteminin farklı kafa pozisyonlarında tutarlı çalışmadığını göstermektedir. Diğer taraftan, seçilen modeller içerisinde en yüksek doğruluk oranlarına sahip olan SCRFD yüz tespit yöntemi, düz kafa pozisyonunda %93, yan kafa pozisyonda ise %92 doğruluk oranı ile yüz tespiti yapmıştır. İlgili yüz tespit yöntemi, hem farklı kafa pozisyonlarında vermiş olduğu istikrarlı sonuçlar hem de sağlamış olduğu yüksek doğruluk oranları sebebiyle yakın kızılötesi görüntüler üzerinde çalışabilecek başarılı ve tutarlı bir yüz tespit yöntemi olmuştur. Seçilen modeller arasındaki en güncel yüz tespit yöntemi olan YuNet ise, en düşük doğruluk oranlarına sahip olmuş ve yakın kızılötesi görüntüler üzerinde yeterli tespit başarısı gösterememiştir.

Tablo 3.1. Yüz tespit yöntemlerinin düz kafa pozisyonunda karşılaştırılması

Yüz Tespit Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
BlazeFace	0.7594	0.7568	0.8734	0.8050	0.6792
CenterFace	0.8886	0.8872	0.9206	0.8979	0.8193
Dlib (HOG)	0.7989	0.8697	0.8017	0.8293	0.7140
DSFD	0.9342	**0.9241**	0.9537	0.9376	0.8863
FaceBoxes	0.8759	0.8591	0.9208	0.8851	0.8115
RetinaFace	0.9238	0.9096	0.9500	0.9283	0.8699
RetinaFace_MobileNet	0.8921	0.8794	0.9245	0.8990	0.8326
SCRFD	**0.9367**	0.9215	**0.9609**	**0.9400**	**0.8897**
YuNet	0.5283	0.8871	0.4661	0.6071	0.4632

Tablo 3.2. Yüz tespit yöntemlerinin yan kafa pozisyonunda karşılaştırılması

Yüz Tespit Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
BlazeFace	0.5913	0.6126	0.8435	0.6979	0.5550
CenterFace	0.8495	0.8205	0.9366	0.8658	0.7690
Dlib (HOG)	0.5417	0.6038	0.5984	0.5969	0.4858
DSFD	0.9162	0.9183	0.9313	0.9224	0.8589
FaceBoxes	0.8206	0.7861	0.9244	0.8417	0.7436
RetinaFace	0.9202	**0.9237**	0.9335	0.9258	0.8647
RetinaFace_MobileNet	0.6430	0.6939	0.7575	0.7199	0.6198
SCRFD	**0.9269**	0.9154	**0.9525**	**0.9316**	**0.8748**
YuNet	0.3666	0.5670	0.3425	0.4217	0.3290

Tablo 3.3. Yüz tespit yöntemlerinin bütün kafa pozisyonlarında karşılaştırılması

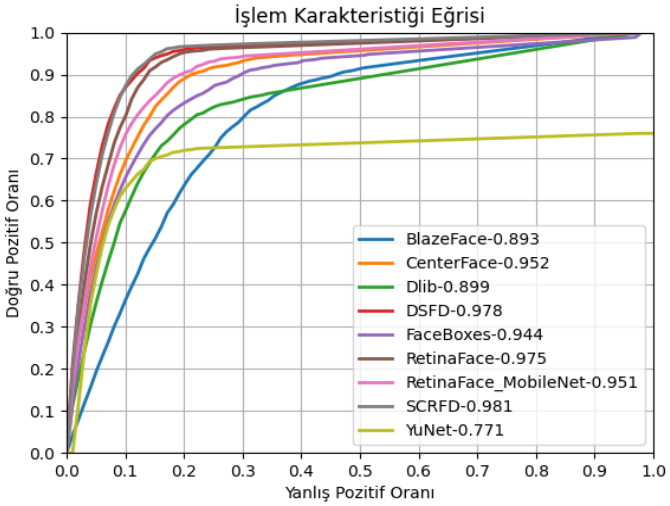
Yüz Tespit Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
BlazeFace	0.6162	0.6217	0.8854	0.7164	0.5757
CenterFace	0.8626	0.8428	0.9312	0.8766	0.7858
Dlib (HOG)	0.5041	0.5915	0.6195	0.5973	0.4875
DSFD	0.9223	**0.9202**	0.9388	0.9275	0.8681
FaceBoxes	0.8391	0.8105	0.9232	0.8562	0.7663
RetinaFace	0.9214	0.9190	0.9390	0.9266	0.8664
RetinaFace_MobileNet	0.6599	0.6937	0.8017	0.7370	0.6376
SCRFD	**0.9302**	0.9174	**0.9553**	**0.9344**	**0.8798**
YuNet	0.4207	0.6741	0.3838	0.4837	0.3739

Yüz tespiti derin sinir ağı modellerin performanslarının değerlendirilmesi için bir diğer karşılaştırmalı analiz İşlem Karakteristiği Eğrisi (Receiver Operating Characteristic Curve, ROC Curve) kullanılarak yapılmıştır. ROC, genellikle sınıflandırma modellerinin performansını değerlendirmek için kullanılan bir eğridir. Bu eğri, ilgili modelin farklı eşik noktalarındaki performansının grafiksel bir temsilidir. Bu temsil, modelin ürettiği yanlış pozitif oranları (false positive rate) ve doğru pozitif oranlarının (true positive rate) değerleri ile oluşturulmaktadır. ROC eğrisinin altındaki alan (area under the ROC Curve, AUC) ise, uygulanan testin genel doğruluğunun bir ölçüsü olarak kullanılmaktadır. Daha yüksek bir AUC değeri, gerçekleştirilen tespit işlemlerimlerinin daha başarılı olduğu anlamına gelmektedir.

Kullanılan veri seti içerisinde düz ve yan kafa pozisyonlarına göre ayrıştırılan yakın kızılötesi görüntüler ayrı ayrı ele alınmış ve her bir durum için bir ROC eğrisi oluşturulmuştur. Düz kafa pozisyonunda alınmış görüntülerin kullanılması ile oluşturulan

ROC eğrisi Şekil 3.7'de, yan kafa pozisyondaki görüntüler ile oluşturulan ROC eğrisi Şekil 3.8'de ve bütün görüntülerin girdi olarak verildiği ROC eğrisi ise Şekil 3.9'da yer almaktadır.

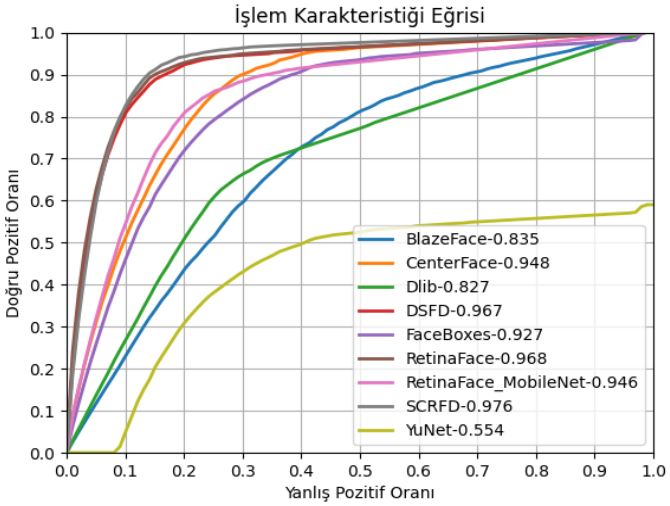
Şekil 3.7'de bulunan düz kafa pozisyonundaki yakın kızılötesi görüntülerin kullanımı ile oluşturulan ROC eğrisinde, en yüksek AUC değerine sahip olan model SCRFD olmuştur. Tablo 3.1'de yer alan diğer performans değerlendirme metriklerinde elde edilen değerlere göre ROC eğrisinden yaklaşık %5 oranında daha yüksek değerler alındığı görülmektedir. Bu durum diğer başarılı sonuçlar üreten RetinaFace ve DSFD modelleri için de aynı şekildedir. Diğer taraftan, başarı sıralamasında ortalarda yer alan CenterFace, FaceBoxes ve RetinaFace_MobileNet modellerinde bu fark yaklaşık %7 oranlarındadır. En düşük başarının elde edildiği BlazeFace ve Dlib'de yaklaşık %11, YuNet modelinde ise yaklaşık %20 oranında bir fark gözlenmiştir. Bu fark değerleri, ilgili modellerin kararlılıkları ve stabilitelerini ortaya koymaktadır. Önceki yapılan analizleri destekler şekilde, en kararlı sonuçlar üreten modeller sırasıyla SCRFD, DSFD ve RetinaFace olurken; en kararsız sonuç üreten model YuNet olmuştur.



Şekil 3.7. Düz kafa pozisyonunda İşlem Karakteristiği Eğrisi

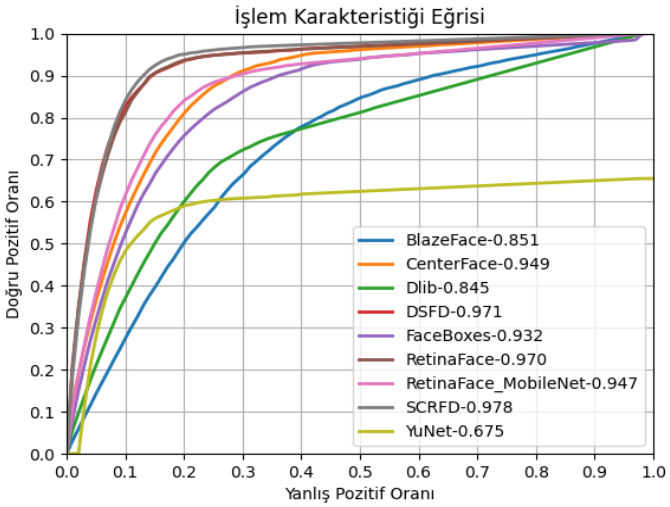
Şekil 3.8'de verilen yan kafa pozisyonundaki yakın kızılötesi görüntülerin kullanımı ile oluşturulan ROC eğrisinde, en başarılı model 0,976 AUC değeri ile SCRFD, en az başarılı model ise 0,554 AUC değeri ile YuNet olmuştur. AUC değerinde 0,5 değeri, yüz bölgesi ile diğer bölgeler arasında ayrımın neredeyse olmadığı anlamına gelmektedir.

Bu sebeple, YuNet tarafından yakın kızılötesi görüntüler üzerinde tespit edilen yüz bölgelerinin oldukça başarısız anlaşılmaktadır. Bu durum, YuNet modelinin üretmiş olduğu yanlış tespit noktalarından kaynaklanmaktadır. Şekil 3.6 (l)'de yer alan görsel incelendiğinde, eğrinin üretmiş olduğu değer ilgili görseli destekler niteliktedir. Diğer taraftan, bütün modellerin tespit ettikleri yüz bölgelerinin, yan kafa pozisyonunda daha az başarılı olduğu görülmektedir. Verilen ROC eğrisindeki bu durum sadece yan kafa pozisyonundaki kızılötesi görüntülerin analiz edildiği Tablo 3.2'de ve bütün görüntülerin yer aldığı Tablo 3.3'te hesaplanan değerler ile örtüşmektedir.



Şekil 3.8. Yan kafa pozisyonunda İşlem Karakteristiği Eğrisi

Şekil 3.9 , hem düz hem de yan kafa pozisyonlarında alınmış yakın kızılötesi görüntülerin beraber ele alındığı durumda üretilmiş ROC eğrisini göstermektedir. İlgili ROC eğrisi incelendiğinde; 0,978 AUC değeri ile yine en başarılı model SCRFD olmuştur. Diğer analizleri destekler nitelikte elde edilen bu sonuçlarla, yüz tespiti derin sinir ağı modellerinin yakın kızılötesi görüntülerde sayısal karşılaştırması kapsamında en başarılı modelin SCRFD modeli olduğu görülmektedir. Diğer taraftan YuNet, 0.675 AUC değeri ile karşılaştırılan modeller arasında en düşük başarı oranına sahip model olmuştur. Diğer modellerin başarı değerleri ise, yan ve düz kafa pozisyonlarında elde edilen değerlerin yaklaşık ortalaması olacak şekilde bulunmuştur. Modellerin başarı sıralamaları, Tablo 3.3'te verilen diğer performans değerleri ile aynı sıralamada olacak şekilde çıkmıştır.

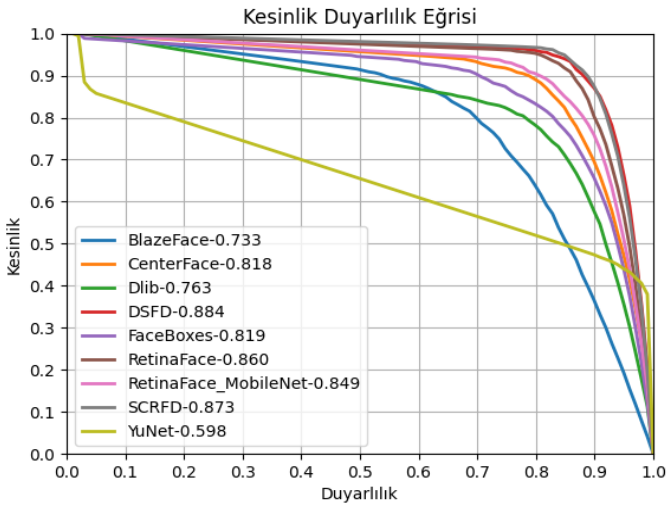


Şekil 3.9. Bütün kafa pozisyonlarında İşlem Karakteristiği Eğrisi

Kesinlik Duyarlılık Eğrisi (Precision Recall Curve, PR Curve); ikili sınıflandırma görevlerinde özellikle dengesiz veri setlerinin bulunduğu senaryolarda kullanılan bir değerlendirme ölçütüdür. Bu ölçüt, farklı sınıflandırma eşikleri için kesinlik ve duyarlılık değerleri arasındaki dengeye ilişkin bilgi sağlayan bir eğri üretmektedir. Bu eğride kesinlik değerinin yüksek olması, modelin yanlış pozitif oranının düşük olup pozitif vakaları doğru tanımlamada iyi olduğunu göstermektedir. Duyarlılık değerinin yüksek olması ise, modelin yanlış negatif oranının düşük olduğunu göstermekte olup pozitif vakaları tespit etmede modelin iyi olduğu anlamına gelmektedir. Bu çıkarım doğrultusunda, yakın kızılötesi görüntüler üzerinde çalıştırılan derin sinir ağı yüz tespit modellerinin PR Eğrisi ile karşılaştırılması yapılmıştır. Bu bağlamda, düz kafa pozisyonu, yan kafa pozisyonu ve bütün kafa pozisyonlarında ele alınan görüntüler ayrı ayrı analiz edilmiştir. Düz kafa pozisyonunda alınan görüntüler kullanılarak oluşturulan PR Eğrisi Şekil 3.10'da, yan kafa pozisyonunda alınan görüntülerin kullanımı ile oluşturulan PR Eğrisi Şekil 3.11'de ve bütün görüntüler ile oluşturulan PR Eğrisi ise Şekil 3.12'de yer almaktadır. Kesinlik duyarlılık eğrisinin altındaki alan (the area under the precision recall curve, AUC-PR), modelin genel performansını özetlemek için kullanılan bir değerdir. Bu değerler, ilgili şekillerde yer alan grafiklerde verilmiştir.

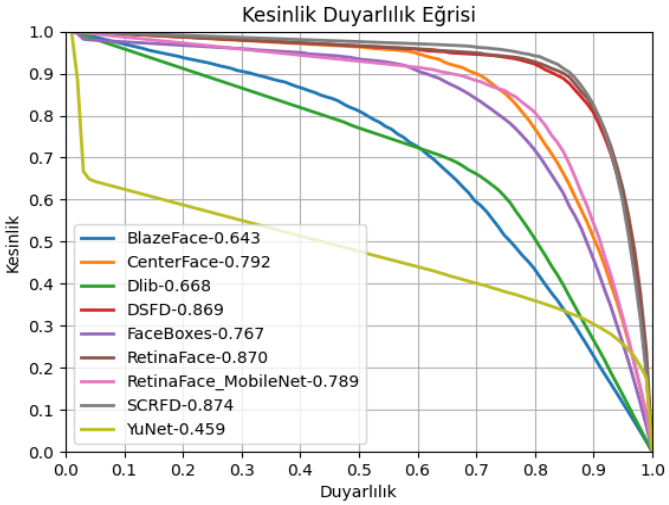
Düz kafa pozisyonundaki yakın kızılötesi görüntüler ile oluşturulan Şekil 3.10'daki PR Eğrisi incelendiğinde, en yüksek AUC-PR değeri 0,884 ile DSFD modeli tarafından

elde edilmiştir. En başarılı ikinci model ise 0,873 AUC-PR değeri ile SCRFD olmuştur. En düşük AUC-PR değeri ise, 0,598 ile YuNet tarafından çıkarılmıştır. Hem elde etmiş olduğu değer hem de ilgili eğrideki kıvrımlar, YuNet modelinin yakın kızılötesi görüntülerde istikrarlı bir model olmadığını ve kararsız yüz tespiti işlemleri gerçekleştirdiğini göstermektedir. Diğer taraftan, kullanılan modellerin PR Eğrisindeki performans değerleri ile Tablo 3.1'de verilmiş kesinlik ve duyarlılık değerlerinin sonuçları arasında benzerik olduğu görülmüştür.



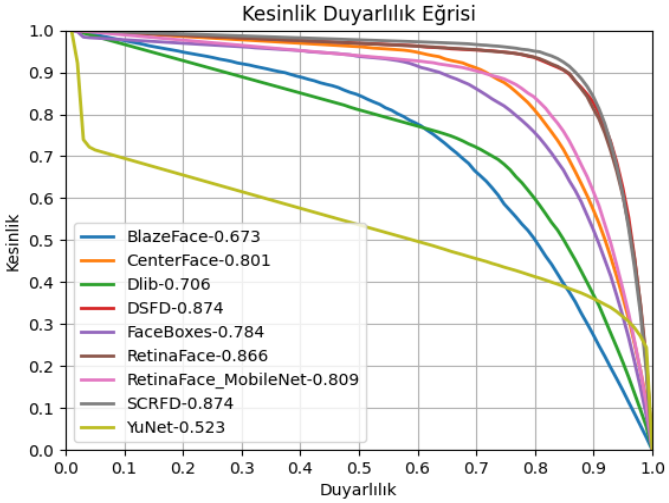
Şekil 3.10. Düz kafa pozisyonunda Kesinlik Duyarlılık Eğrisi

Şekil 3.11'de, yan kafa pozisyonundaki yakın kızılötesi görüntülerin kullanımı ile oluşturulmuş PR Eğrisi verilmiştir. Diğer gerçekleştirilen analizlerde olduğu gibi burada da bütün modellerin başarılarının düştüğü görülmektedir. Bu durum, yan kafa pozisyonunda gerçekleştirilen yüz tespitinde, kullanılan modellerin daha fazla hata payına sahip olduklarını göstermektedir. Verilen PR Eğirisinde en yüksek AUC-PR değeri, 0,874 değeri ile SCRFD modeline aittir. En düşük AUC-PR değeri ise 0,459 ile YuNet tarafından elde edilmiştir. YuNet modelinin 0,5'ten daha düşük bir değer ile sonuç üretiyor olması, ilgili modelin yakın kızılötesi görüntüler üzerinde yan kafa pozisyonda yapılan tespitlerde başarısız olduğu anlamına gelmektedir.



Şekil 3.11. Yan kafa pozisyonunda Kesinlik Duyarlılık Eğrisi

Şekil 3.12'de, bütün kafa pozisyonundaki yakın kızılötesi görüntülerin kullanımı ile elde edilmiş PR Eğrisi yer almaktadır. Bu eğride, hem yan hem de düz kafa pozisyonunda alınmış yakın kızılötesi görüntüler kullanılmıştır. İlgili eğride, DSFD ve SCRFD modelleri, 0,874 AUC-PR değeri üreterek en yüksek başarıya sahip modeller olmuştur. YuNet modeli ise 0,523 AUC-PR değeri ile diğer senaryolarda oluduğu gibi bu durumda da en düşük başarıyı gösteren model olmuştur. Verilen eğride üretilen değerler Tablo 3.3'te bulunmuş değerleri destekler niteliktedir. Bulunan değerler doğrultusunda belirlenen modellerin başarı sıralamaları bütün çalışmalarda benzerlik göstermektedir.



Şekil 3.12. Bütün kafa pozisyonlarında Kesinlik Duyarlılık Eğrisi

3.3.3. Yüz Tespiti Modellerinin İşlem Zamanlarının Karşılaştırması

Yüz tespiti modellerinin işlem zamanlarının karşılaştırması, farklı modellerin aynı işlemi gerçekleştirmek için tükettiği süreleri ve performanslarını değerlendirmek amacıyla yapılmıştır. Yapılan çalışma, hem CPU hem de GPU üzerinde yürütülerek, Tablo 3.4'te görüleceği üzere ayrı ayrı analiz sonuçları elde edilmiştir. Elde edilen sonuçlara göre; CPU üzerinde 0,011 saniyelik çalışma süresi ile BlazeFace, GPU üzerinde ise 0,004 saniyelik çalışma süresi ile FaceBoxes en hızlı model olmuştur. Diğer taraftan; 3 saniyenin üzerinde işlem zamanı gerektiren Dlib, CPU üzerinde en yavaş çalışan model olurken, RetinaFace 0,8 saniye işlem zamanı ile GPU üzerinde en yavaş çalışan model olmuştur.

Tablo 3.4. İşlem zamanı karşılaştırma tablosu

Yüz Tespit Yöntemi	İşlem Zamanı (sn)	
	CPU	GPU
BlazeFace	**0,011**	0,006
CenterFace	0,460	0,170
Dlib (HOG)	3,010	-
DSFD	2,300	0,260
FaceBoxes	0,100	**0,004**
RetinaFace	1,800	0,800
RetinaFace_MobileNet	0,430	0,058
SCRFD	0,055	0,047
YuNet	0,410	-

Gerçekleştirilen değerlendirme sürecinde; işlem zamanı karşılaştırması ile, ilgili modellerin kullandığı omurga ağ, sahip olduğu parametre sayısı ve kayan nokta operasyonu (Floating Point Operations, FLOP) miktarları da göz önüne alınmıştır. Bahsedilen çalışmaya ilişkin karşılaştırma verileri Tablo 3.5'de yer almaktadır. Sonuç olarak; Tablo 3.4 ve Tablo 3.5 beraber incelendiğinde, modellerin sahip oldukları parametre sayıları ve FLOP miktarları ile işlem zamanları arasında ilişki olduğu görülmektedir.

Tablo 3.5. Modellere ait omurga ağı, parametre ve FLOP karşılaştırma tablosu

Yüz Tespit Yöntemi	Omurga Ağı	Param (M)	FLOP (G)
BlazeFace	BlazeNet*	0,12	0.71
CenterFace	MobileNetV2	3,48	0,88
DSFD	ResNet152	120,06	259,55
FaceBoxes	FaceNet**	1,01	0,27
RetinaFace	ResNet50	29,50	37,59
RetinaFace_MobileNet	MobileNet0,25	0,46	0,80
SCRFD	Basic Res	0,82	2,50
YuNet	MobileNet	0,75	1,49

* BlazeNet -> iyileştirilmiş MobileNet
** FaceNet -> RDCL ve MSCL

3.4. Sınırlayıcı Kutu Belirleme Yöntemlerinin Karşılaştırılması

Sınırlayıcı kutu belirleme yöntemleri (bounding box ensemble), nesne tespitinde elde edilen sınırlayıcı kutunun doğruluğunu ve hassasiyetini geliştirmek için birden fazla modeli veya algoritmayı birleştiren teknikleri ifade etmektedir. Bu yöntemler, farklı modellerin güçlü yönlerinden yararlanarak daha doğru sınırlayıcı kutu tahminleri üretmeyi amaçlamaktadır. Bu bölümde, bu amaç doğrultusunda geliştirilmiş NMS, NMW ve WBF sınırlayıcı kutu belirleme yöntemleri, basit olarak kullanılabilecek OA yöntemi ve bu çalışma kapsamında önerilen NMA ve KO yöntemleri karşılaştırılmaktadır. Yapılan analizler kapsamında; performans değerlendirme metriği olarak Doğruluk, Kesinlik, Duyarlılık, F1 Skor ve Jaccard kullanılmıştır.

Bu bölümde gerçekleştirilen analizlerde, Bölüm 3.3'de yapılan deneylerde en iyi sonucu veren yüz tespiti derin sinir ağı modelleri kullanılmıştır. Bu bağlamda, en iyi sonucu veren modeller beşli, dörtlü, üçlü ve ikili olacak şekilde gruplandırılarak ürettikleri sınırlayıcı kutular birleştirilmiştir. Bu sayede, mevcut ve önerilen yöntemlerin kullanılan model sayısına hassasiyeti ve performans oranlarındaki değişimi analiz edilmiştir.

3.4.1. Yöntemlerin Görsel Karşılaştırması

Sınırlayıcı kutu belirleme yöntemlerinin farklı senaryolar altında performanslarının analiz edilmesi, ilgili yöntemlerinin kullanım verimliliğini artırıcı niteliktedir. Bu sebeple, ilgili modellerin üretmiş oldukları sınırlayıcı kutular; beşli, dörtlü, üçlü ve ikili olacak şekilde

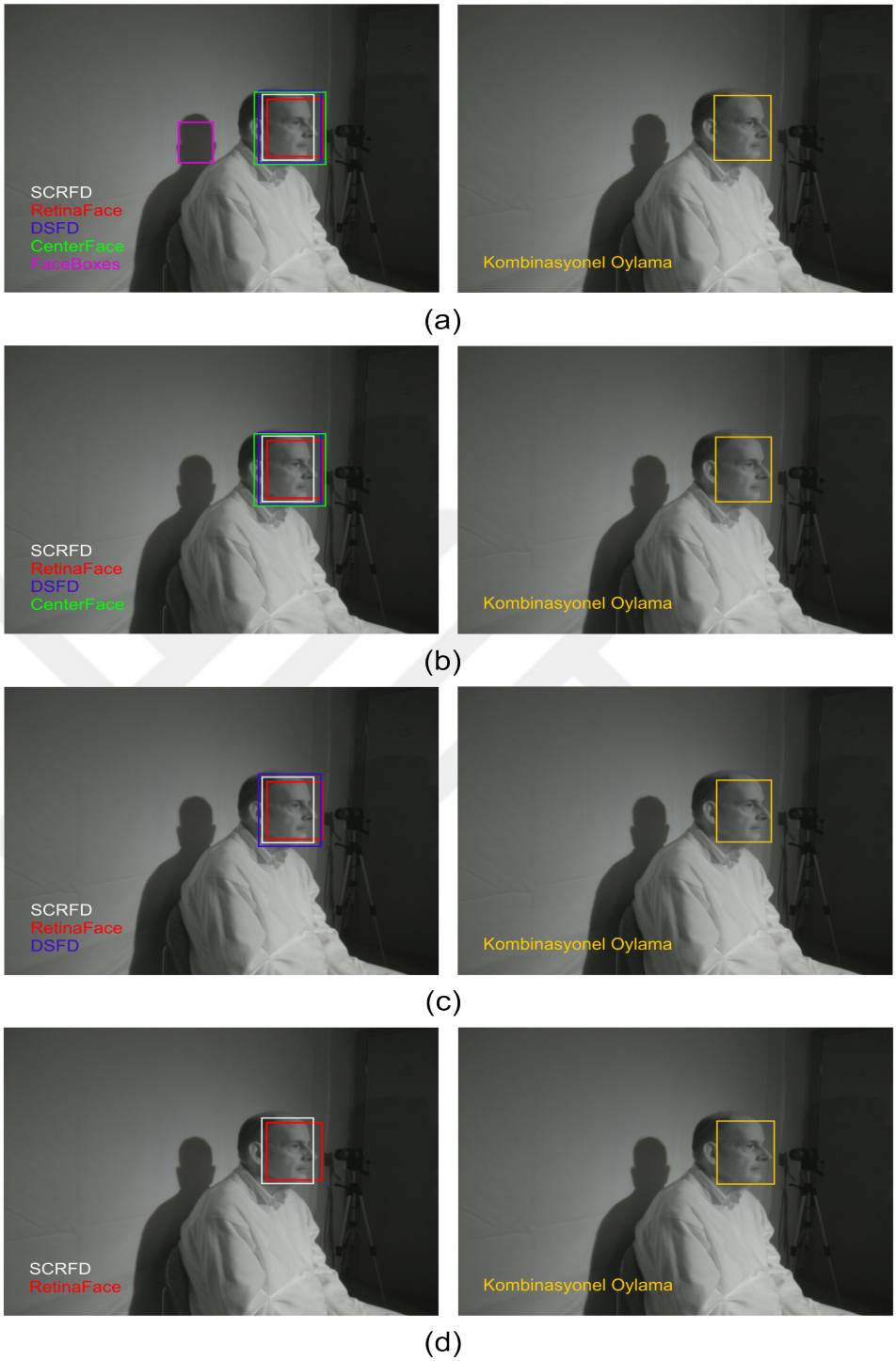
analiz edilmiştir. Bu bağlamda, Kombinasyonel Oylama yöntemin farklı senaryolarda üretmiş olduğu sonuç sınırlayıcı kutular Şekil 3.13'te, Maksimum Olmayan Ortalama yönteminin üretmiş olduğu sınırlayıcı kutular ise Şekil 3.14'te yer alan görsellerde her bir girdi senaryo için ayrı ayrı yer almaktadır.

Şekil 3.13 (a) incelendiğinde görüleceği üzere, verilen girdi görüntü için FaceBoxes modeli yanlış yüz tespiti yapmakta, CenterFace ise olması gereken yüz bölgesinden daha geniş bir alanı işaretlemektedir. Ayrıca diğer modellerden daha başarılı sonuçlar üretmesine rağmen SCRFD modelinin çıktısı olan sınırlayıcı kutuda kişinin burun bölgesi tam olarak alınmamaktadır. Diğer taraftan, Kombinasyonel Oylama yönteminin beş farklı modelin ürettiği sınırlayıcı kutuları kullanarak tüm modellerden daha başarılı bir sınırlayıcı kutu ürettiği görülmektedir.

Şekil 3.13 (b)'de en iyi sonucu veren dört model ele alınmıştır. FaceBoxes'ın kullanılmadığı bu senaryoda, beş modelin kullanıldığı senaryo ile aynı yüz bölgesinin tespit edildiği görülmektedir. Bu durum, Kombinasyonel Oylama yönteminin kullanmakta olduğu eliminasyon aşamalarının alakasız yüz bölgesi tespit eden FaceBoxes'ın üretmiş olduğu sınırlayıcı kutuyu elemiş olduğunu göstermektedir.

Şekil 3.13 (c) en iyi sonucu veren üç modelin kullanıldığı senaryoya aittir. Bu senaryoda elde edilen sonuçta, yüz bölgesi alanının önceki senaryolara göre daha başarılı tespit edildiği görülmektedir. Beş ve dört modelin kullanıldığı senaryolarda, yüz bölgesinin dışına çıkılıp baş bölgesi de yüz alanına dahil edildiği görülmektedir. Üç modelin kullanıldığı senaryoda, baş bölgesinden uzaklaşma olmuş olup tespiti amaçlanan yüz bölgesi daha başarılı tespit edilmiştir.

Şekil 3.13 (d) en yüksek doğruluk oranlarına sahip iki modelin girdisi ile Kombinasyonel Oylama yöntemi tarafından üretilen yeni sınırlayıcı kutuyu göstermektedir. Bu senaryoda, kullanılan modeller arasında en başarılı sonuç veren SCRFD modeli tarafından üretilen sınırlayıcı kutunun baş, çene ve kulak bölgelerinden taşmalara sahip olduğu görülmektedir. Üretilen yeni sınırlayıcı kutuda ise bu bölgeleri sınırlayan alanların yüz bölgesine daha yakın olduğu görülmektedir.



Şekil 3.13. En iyi sonucu veren modellerin kullanımı ile Kombinasyonel Oylama yönteminden yeni sınırlayıcı kutu üretilmesi: (a) beş model, (b) dört model, (c) üç model, (d) iki model

Şekil 3.14 (a), kullanılan modeller arasında en iyi sonucu veren beş modelin NMA yöntemine girdi olarak verilmesi sonucu üretilen yeni sınırlayıcı kutuyu göstermektedir. Görselde görüleceği üzere, önerilen yöntem yüz alanından daha geniş bir alanı kaplayan

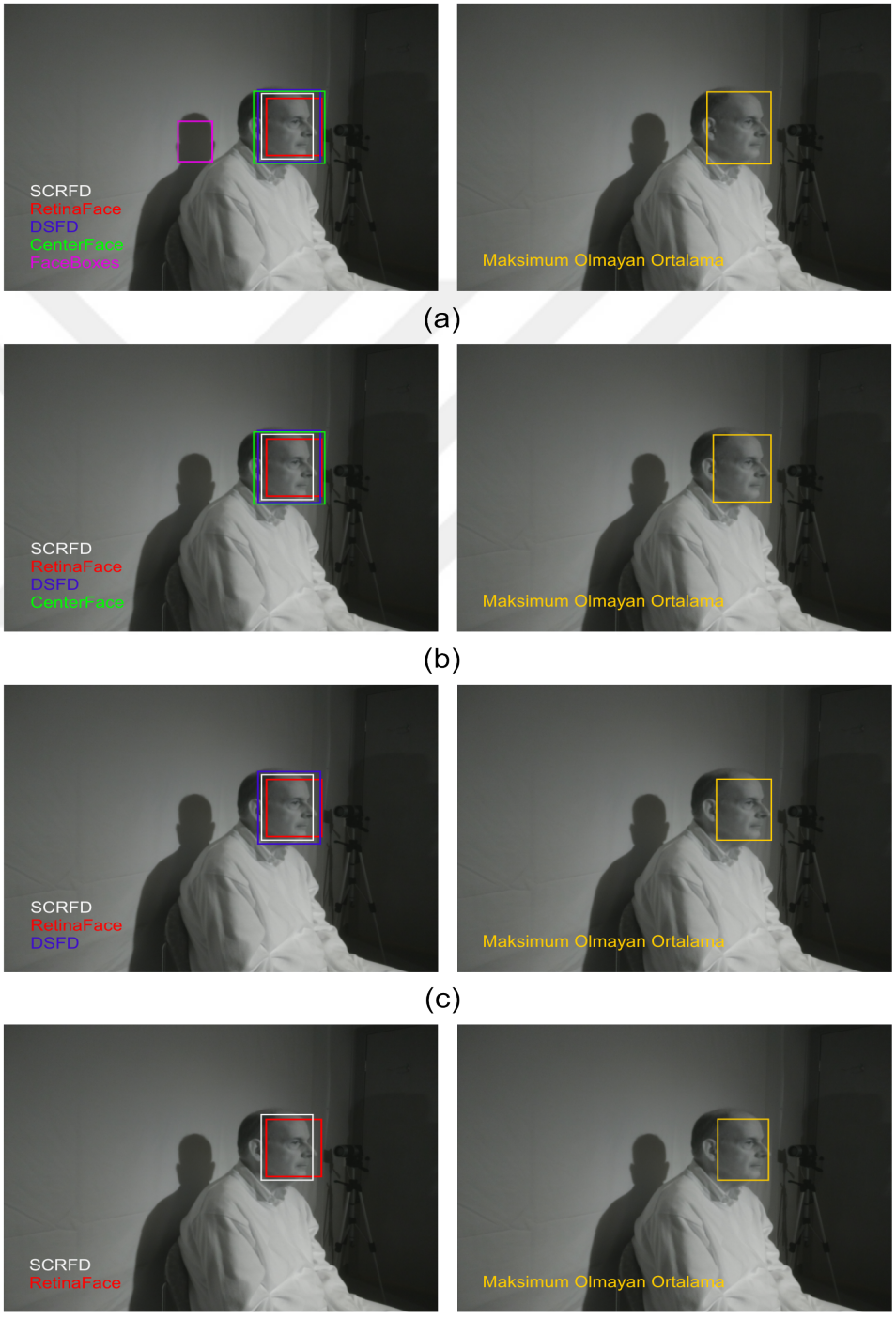
yeni bir sınırlayıcı kutu çizilmiştir. Bu durum, CenterFace modelinin en ve boy olarak, DSFD modelinin ise boy olarak geniş bir konumlandırma yapmasından ilgili yöntemin etkilendiğini göstermektedir. Ayrıca FaceBoxes modelinin yapmış olduğu yanlış yüz bölgesi tespiti, önerilen yöntemin yüz konumlandırmasını etkileyen diğer önemli bir faktör olmuştur. Tüm bu durumlara rağmen, NMA yönteminin yüz bölgesinin tamamını kapsayan bir konumlandırma yapmış olduğu görülmektedir. Olması gerekenden daha geniş bir alanı işaretliyor olması ise ilgili yöntemin geliştirmeye açık yönünü göstermektedir.

Şekil 3.14 (b), en iyi sonucu veren dört yüz tespiti derin sinir ağı modeli kullanılarak elde edilmiştir. Bu senaryoda, modeller arasında en düşük başarıya sahip olan FaceBoxes modeli çıkarılmıştır. Yanlış yüz bölgesi tespiti yapan FaceBoxes modelinin kullanılmamasının, NMA yönteminin üretmiş olduğu yeni sınırlayıcı kutunun konumlandırmasını iyileştirdiği verilen görselde görülmektedir. Beşli senaryoda neredeyse tamamı dahil edilmiş olunan kulak bölgesi bu durumda çok daha az işaretlenmiştir. Ayrıca çizilen sınırlayıcı kutu alanının boydan daraldığı ve çene bölgesi altı fazla çizilen bölgenin kapandığı görülmektedir.

Şekil 3.14 (c)'de en iyi sonucu veren üç modelin kullanıldığı senaryoya ait görsel bulunmaktadır. Bu senaryoda, yüz bölgesi alanının önceki senaryolara göre daha başarılı tespit edildiği görülmektedir. Önceki senaryolarda işaretlenen kulak bölgesi üçlü durumda işaretlenmemiştir. Çene bölgesinin altındaki boğaz alanı ve yüzün dışında kalan kafa bölgesi yine işaretlenen sınırlayıcı kutu alanında bulunmamaktadır. Görülen bu iyileşmeler, gerçekleştirilen senaryonun sayısal performans verilerini destekler niteliktedir. Her bir senaryoya ait performans değerlendirmelerinin yer aldığı Tablo 3.6 - 3.9 arasındaki bütün performans verileri arasından üçlü senaryoda NMA yönteminin daha başarılı olduğu görülmektedir.

Şekil 3.13 (d)'de en yüksek doğruluk oranlarına sahip iki modelin kullanıldığı durumda NMA yöntemi tarafından üretilen yeni sınırlayıcı kutu yer almaktadır. Üretilen sınırlayıcı kutunun işaretlediği alanlarda burun bölgesini tam olarak alamadığı görülmektedir. Bu durum, kullanılan modeller arasında en yüksek performans değerlerinin alındığı SCRFD modelinin burun bölgesini tam alamaması, RetinaFace modelinde ise ufak taşmaların

olması sebebiyle oluşmuştur. Ancak üretilen yeni sınırlayıcı kutuda NMA yönteminin; kulak, çene ve baş bölgelerinde ideale yakın bir konumlandırma yaptığı ve genel olarak her iki girdi modelden de daha başarılı bir konumlandırma yaptığı ilgili şekilden çıkarılabilmektedir.



Şekil 3.14. En iyi sonucu veren modellerin kullanımı ile Maksimum Olmayan Ortalama yönteminden yeni sınırlayıcı kutu üretilmesi: (a) beş model, (b) dört model, (c) üç model, (d) iki model

Karşılaştırmalar neticesinde; sınırlayıcı kutu belirleme yöntemlerinin birçoğunun mevcut yüz tespit yöntemlerinin başarısından daha yüksek sonuçlar ürettiği görülmektedir. Bu durum sınırlayıcı kutu belirleme yöntemlerinin mevcut yüz veya nesne algılayıcılarının tespit doğruluğunu artırıcı nitelikte olduğunu kanıtlamaktadır.

3.4.2. Yöntemlerin Sayısal Karşılaştırması

Gerçekleştirilen çalışmaya ilişkin sonuç değerler Tablo 3.6 - 3.9'da yer almaktadır. Sayısal sonuçlar incelendiğinde; en iyi beş modelin (Tablo 3.6) kullanıldığı senaryoda; önerilen KO yöntemi Kesinlik haricindeki bütün performans değerlendirme metriklerinde en başarılı sonucu üretmiştir. En iyi dört modelin (Tablo 3.7) kullanıldığı senaryoda, Doğruluk ve Duyarlılık metriklerinde, en iyi üç modelin (Tablo 3.8) kullanımda Doğruluk ve Jaccard metriklerinde, en iyi iki modelin (Tablo 3.9) yer aldığı senaryoda ise Duyarlılık ve F1 Skor performans değerlendirme metriklerinde KO yöntemi daha başarılı olmuştur.

Önerilen KO yöntemi, mevcut sınırlayıcı kutu belirleme yöntemlerinden daha başarılı olduğu gibi değerlendirmeye alınan tüm yüz tespiti modellerinden de daha başarılı sonuçlar üretmiştir. Tablo 3.3 ve Tablo 3.6 beraber incelendiğinde; KO dışındaki sınırlayıcı kutu belirleme yöntemlerinin sadece Kesinlik ve Duyarlılık performans değerlendirme metriklerinde SCRFD'den daha başarılı olduğu, KO yönteminin ise bütün metriklerde girdi olarak verilen modellerden daha başarılı sonuçlar ürettiği görülmektedir.

Tablo 3.6. En iyi sonucu veren beş modelin (SCRFD, RetinaFace, DSFD, CenterFace, FaceBoxes) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınılayıcı Kutu Belirleme Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
NMS	0.9232	**0.9219**	0.9389	0.9284	0.8694
NMW	0.9241	0.9068	0.9557	0.9288	0.8702
WBF	0.9221	0.9027	0.9567	0.9270	0.8671
NMA	0.9243	0.9019	0.9564	0.9268	0.8586
OA	0.9041	0.8827	0.9447	0.9105	0.8490
KO	**0.9326**	0.9205	**0.9591**	**0.9374**	**0.8801**

Tablo 3.7. En iyi sonucu veren dört modelin (SCRFD, RetinaFace, DSFD, CenterFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınılayıcı Kutu Belirleme Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
NMS	0.9231	**0.9218**	0.9389	0.9283	0.8692
NMW	0.9278	0.9162	0.9525	0.9323	0.8761
WBF	0.9271	0.9140	0.9539	0.9315	0.8748
NMA	0.9159	0.8989	0.9501	0.9216	0.8582
OA	0.9250	0.9126	0.9602	**0.9557**	**0.8771**
KO	**0.9337**	0.9066	**0.9689**	0.9348	0.8749

Tablo 3.8. En iyi sonucu veren üç modelin (SCRFD, RetinaFace, DSFD) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınılayıcı Kutu Belirleme Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
NMS	0.9231	0.9218	0.9389	0.9283	0.8692
NMW	0.9264	0.9205	0.9453	0.9310	0.8738
WBF	0.9271	0.9204	0.9454	0.93105	0.8748
NMA	0.9283	**0.9304**	0.9402	0.9283	0.8694
OA	0.9268	0.9203	**0.9461**	**0.9364**	0.8746
KO	**0.9366**	0.9280	0.9455	0.9353	**0.8819**

Tablo 3.9. En iyi sonucu veren iki modelin (SCRFD, RetinaFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınılayıcı Kutu Belirleme Yöntemi	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
NMS	0.9216	0.9188	0.9396	0.9268	0.8666
NMW	0.9266	0.9192	0.9475	0.9312	0.8742
WBF	0.9268	0.9192	0.9478	0.9314	0.8745
NMA	0.9273	**0.9239**	0.9390	0.9266	0.8664
OA	**0.9345**	0.9123	0.9557	0.9374	**0.8817**
KO	0.9327	0.9104	**0.9701**	**0.9376**	0.8774

4. BÖLÜM

TARTIŞMA, SONUÇ VE ÖNERİLER

Sunulan yüksek lisans tezi kapsamında yapılan çalışmalar ve deneyler, Bölüm 4.1'de ele alınmaktadır. Bu bağlamda gerçekleştirilen çalışmalar hakkında bilgi verilmektedir. Gerçekleştirilen çalışmalar neticesinde elde edilen sonuçlar, bu sonuçların değerlendirmesi ve getirilen öneriler ise Bölüm 4.2'de açıklanmaktadır.

4.1. Tartışma

Yetersiz aydınlatma şartlarında başarılı şekilde yüz tespiti yapılabilmesi için yakın kızılötesi görüntülerden yararlanılabilmektedir. Bu çalışmada literatürde yaygın olarak kullanılan yüz tespiti yöntemleri yakın kızılötesi görüntüler üzerinde karşılaştırılmıştır. Yürütülen çalışmada; literatürde etiketli yakın kızılötesi yüz görüntülerinden oluşan bir veri seti bulunmaması sebebiyle TUFTS yüz görüntüsü veri seti etiketlenerek, yakın kızılötesi görüntüler üzerinde yüz tespitleri yapılabilmesi için referans değerler oluşturulmuştur. Elde edilen referans değerlerin ve beş farklı performans değerlendirme metriğinin kullanılması ile dokuz farklı yüz tespit yönteminin yakın kızılötesi görüntüler üzerinde karşılaştırmalı analizi yapılmıştır. Analiz kapsamında; düz kafa pozisyonu, yan kafa pozisyonu ve bütün kafa pozisyonları olmak üzere ilgili veriler ayrı ayrı ele alınmıştır. Bu sayede, yüz tespit yöntemlerinin yakın kızılötesi görüntülerde istikrarlı çalışma durumları ve tutarlılıkları ortaya çıkarılmıştır. Bunlara ilaveten, ilgili yüz tespit yöntemlerinin farklı donanım bileşenleri üzerinde işlem zamanı karşılaştırmaları yapılarak performansları değerlendirilmiştir. Ayrıca, yüz tespiti modellerinin başarımlarını artırmak için iki yeni sınırlayıcı kutu belirleme yöntemi önerilmiştir. Bu yöntemler yüz tespiti modellerinin ürettiği sınırlayıcı kutu topluluklarını kullanarak daha başarılı yeni bir sınırlayıcı kutu üretmektedir. Bu doğrultuda, önerilen

sınırlayıcı kutu belirleme yöntemleri literatürde kullanılan dört sınırlayıcı kutu belirleme yöntemi ile karşılaştırılmıştır.

4.2. Sonuç ve Öneriler

Yüz tespit yöntemlerinin farklı kafa pozisyonlarında incelendiği çalışmada, düz kafa pozisyonunda yapılan yüz tespiti işlemlerinin daha başarılı performans verdiği ve yan kafa pozisyonlarında bu başarının azaldığı görülmüştür. Bu durum, genel sonuca bakıldığında kafa pozisyonundan daha az etkilenen yüz tespit yöntemlerinin daha yüksek başarıya sahip olması ile örtüşmektedir. Bütün kafa pozisyonlarında en yüksek doğruluk oranları ile tespit yapan SCRFD, yakın kızılötesi görüntüler üzerinde en başarılı yüz tespiti işlemi gerçekleştiren model olmuştur. RetinaFace ve DSFD modelleri ise sırasıyla en başarılı ikinci ve üçüncü model konumundadır. Seçilen modeller arasında en güncel yüz tespit yöntemi olan YuNet ise, bütün kafa pozisyonlarında en düşük doğruluk oranları ile yüz tespiti gerçekleştirmiştir. Diğer taraftan, Retina MobileNet, BlazeFace veya Dlib düz kafa pozisyonunda daha başarılı sonuçlar üretirken yan kafa pozisyonunda yeterli başarıyı sağlayamamıştır. Bu yüz tespit yöntemleri tutarlı sonuçlar üretmek konusunda başarılı olamamış ve istikrarı sağlayamamışlardır.

İşlem zamanı karşılaştırması çalışmaları kapsamında, bütün modeller CPU ve GPU üzerinde ayrı ayrı çalıştırılmışlardır. CPU üzerinde, en hızlı çalışan yüz tespit yöntemi BlazeFace, en yavaş çalışan ise Dlib olmuştur. GPU üzerinde, en hızlı çalışan FaceBoxes, en yavaş çalışan model ise RetinaFace olmuştur. Burada, RetinaFace yüz tespit yönteminin en yüksek başarı oranına sahip ikinci model olmasına rağmen yüksek işlem zamanı gerektirdiği ortaya çıkmaktadır.

Yüz tespiti yöntemlerinin performans değerlendirmelerine ilaveten, sınırlayıcı kutu belirleme yöntemlerinin analizi yapılmıştır. Bu deneylerde, yüz tespit yöntemlerinin sırasıyla en başarılı beş, dört, üç ve iki tanesi seçilerek ürettikleri sınırlayıcı kutular karşılaştırılan sınırlayıcı kutu belirleme yöntemleri ile iyileştirilmiştir. Bu deneyler neticesinde, en başarılı üç modelin kullanımı ile gerçekleştirilen çalışmalarda daha başarılı sonuçlar elde edildiği görülmüştür. Diğer taraftan; önerilen Kombinasyonel Oylama yöntemi; en iyi beş, dört ve üç modelin seçildiği senaryolarda diğer yöntemlere

nazaran daha başarılı sonuçlar üretmiştir. Deneylerde yer verilen ortalama alma yöntemi diğer yöntemlerin başarısını değerlendirme konusunda belirleyici bir kriter olmuştur. En iyi üç modelin kullanıldığı çalışmada NMS ve NMW yöntemleri, en iyi dört modelin yer aldığı çalışmada ise NMS ve NMA yöntemleri ortalamanın altında kalmıştır. Bu durum anılan yöntemlerin bazı senaryolarda basit ortalama alma yaklaşımının dahi altında kalarak beklenilen performansı sergileyemediklerini ortaya koymaktadır. En başarılı iki modelin kullanımı ile gerçekleştirilen çalışmada ise Ortalama Alma yöntemi diğer senaryolara göre daha başarılı olmuştur. Bu durum, sınırlayıcı kutu belirleme yöntemlerinin en az üç model ile kullanılmasının daha verimli olduğunu göstermektedir. Sonuç olarak, sınırlayıcı kutu belirleme yöntemlerinin kullanılması, yüz tespit yöntemleri tarafından üretilen tespit noktalarının başarısını artıracak niteliktedir. Ayrıca, sunulan yöntemlerden Kombinasyonel Oylama yönteminin, mevcut sınırlayıcı kutu belirleme yöntemlerinden daha yüksek başarıya sahip olduğu görülmüştür.

Gelecekte; yapılan çalışmaların daha geniş veri setleri ile yapılması planlanmaktadır. Bu bağlamda, yeni veri setlerinin oluşturulması veya mevcut veri setinin genişletilmesi söz konusu olabilecektir. Mevcut veri setinin genişletilmesi hususunda, NTHU Sürücü Dikkat Dağınıklığı Veri Seti (National Tsing Hua University Driver Drowsiness Dataset) [48], MR-NIRP Veri Seti (MERL-Rice Near-Infrared Pulse Dataset) [49] ve IIT Delhi Yakın Kızılötesi Yüz Veri Seti (IIT Delhi Near IR Face Database) [50]; ilk araştırmalar doğrultusunda belirlenmiştir. Bunlar ve belirlenecek diğer veri setleri dahilindeki görüntülere başvurulabilecektir.

İlerleyen süreçte, önerilen iki yeni sınırlayıcı kutu belirleme yönteminin sadece yakın kızılötesi görüntüler için değil diğer görüntü türlerinde de başarımlarının araştırılması amaçlanmaktadır. Bu bağlamda, WIDER Face [41], PASCAL Face [51], FDDB (Face Detection Data Set and Benchmark) [52], AFW (Annotated Faces in the Wild) [53] ve AFLW (Annotated Facial Landmarks in the Wild) [54] veri setleri üzerinde analiz yapılması planlanmaktadır. Literatürde popüler olarak kullanılan bu ve benzeri veri setleri ile önerilen sınırlayıcı kutu belirleme yöntemlerinin başarısının bütün görüntü türleri için ortaya koyulması gaye edinilmektedir. Tez çalışması sürecinde elde edilen tecrübe ve çıkarılan sonuçların ileriki çalışmalara ışık tutacağı düşünülmektedir.

KAYNAKLAR

1. Kumar, A., Kaur, A., and Kumar, M., 2019. Face detection techniques: a review, **Artificial Intelligence Review**, **52**:927–948.

2. Panetta, K., Wan, Q., Agaian, S., Rajeev, S., Kamath, S., Rajendran, R., Rao, S.P., Kaszowska, A., Taylor, H.A., Samani, A., et al., 2018. A comprehensive database for benchmarking imaging systems, **IEEE transactions on pattern analysis and machine intelligence**, **42**(3):509–520.

3. Zou, Z., Chen, K., Shi, Z., Guo, Y., and Ye, J., 2023. Object detection in 20 years: A survey, **Proceedings of the IEEE**.

4. King, D.E., 2009. Dlib-ml: A machine learning toolkit, **The Journal of Machine Learning Research**, **10**:1755–1758.

5. Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., and Li, S.Z., 2017. Faceboxes: A CPU real-time face detector with high accuracy, *2017 IEEE International Joint Conference on Biometrics (IJCB)*, 1–9, IEEE.

6. Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J., and Huang, F., 2019. DSFD: dual shot face detector, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5060–5069.

7. Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K., and Grundmann, M., 2019. Blazeface: Sub-millisecond neural face detection on mobile gpus, **arXiv preprint arXiv:1907.05047**.

8. Xu, Y., Yan, W., Yang, G., Luo, J., Li, T., and He, J., 2020. CenterFace: joint face detection and alignment using face as point, **Scientific Programming**, **2020**:1–8.

9. Deng, J., Guo, J., Ververas, E., Kotsia, I., and Zafeiriou, S., 2020. Retinaface: Single-shot multi-level face localisation in the wild, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 5203–5212.

10. Guo, J., Deng, J., Lattas, A., and Zafeiriou, S., 2021. Sample and computation redistribution for efficient face detection, **arXiv preprint arXiv:2105.04714**.

11. Wu, W., Peng, H., and Yu, S., 2023. YuNet: A Tiny Millisecond-level Face Detector, **Machine Intelligence Research**, 1–10.

12. Neubeck, A. and Van Gool, L., 2006. Efficient non-maximum suppression, *18th international conference on pattern recognition (ICPR'06)*, volume 3, 850–855, IEEE.

13. Solovyev, R., Wang, W., and Gabruseva, T., 2021. Weighted boxes fusion: Ensembling boxes from different object detection models, **Image and Vision Computing**, **107**:104117.

14. Ning, C., Zhou, H., Song, Y., and Tang, J., 2017. Inception single shot multibox detector for object detection, *2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 549–554, IEEE.

15. Abdulhussien, D.M. and Saud, L.J., 2023. A Review of Face Detection Methods Based on Feature Approach, **Iraqi Journal Of Computers, Communications, Control And Systems Engineering**, **23**(3):1–9.

16. Cho, S.W., Baek, N.R., Kim, M.C., Koo, J.H., Kim, J.H., and Park, K.R., 2018. Face detection in nighttime images using visible-light camera sensors with two-step faster region-based convolutional neural network, **Sensors**, **18**(9):2995.

17. Li, J., Zhang, D., Zhang, K., Hu, K., and Yang, L., 2017. Real-time face detection during the night, *2017 4th International Conference on Systems and Informatics (ICSAI)*, 582–586, IEEE.

18. Liao, S., Jain, A.K., and Li, S.Z., 2015. A fast and accurate unconstrained face detector, **IEEE transactions on pattern analysis and machine intelligence**, **38**(2):211–223.

19. Wang, W., Wang, X., Yang, W., and Liu, J., 2022. Unsupervised face detection in the dark, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, **45**(1):1250–1266.

20. Ozaki, Y., 2021. Infrared spectroscopy—Mid-infrared, near-infrared, and far-infrared/terahertz spectroscopy, **Analytical Sciences**, **37**(9):1193–1212.

21. Beiser, A., 2003. Concepts of Modern Physics, 6 ed., New York: McGraw Hill.

22. Hansen, M.P. and Malchow, D.S., 2008. Overview of SWIR detectors, cameras, and applications, *Thermosense Xxx*, volume 6939, 94–104, SPIE.

23. Zhang, Y. and Qiao, J., 2021. Near-infrared emitting iridium complexes: Molecular design, photophysical properties, and related applications, **Iscience**, **24**(8).

24. Dash, P., Kisku, D.R., Gupta, P., and Sing, J.K., 2022. Fast face detection using a unified architecture for unconstrained and infrared face images, **Cognitive Systems Research**, **74**:18–38.

25. Gao, Z., Zhang, S., Fang, H., Li, L., and Huang, L., 2021. Multi-modal Image Fusion based Improved Face Detection Algorithm in Poor Lighting Conditions, *2021 4th International Conference on Signal Processing and Machine Learning*, 33–38.

26. Ren, S., He, K., Girshick, R., and Sun, J., 2015. *Faster r-cnn: Towards real-time object detection with region proposal networks*, volume 28.

27. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and Berg, A.C., 2016. Ssd: Single shot multibox detector, *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, 21–37, Springer.

28. Tang, X., Du, D.K., He, Z., and Liu, J., 2018. Pyramidbox: A context-assisted single shot face detector, *Proceedings of the European conference on computer vision (ECCV)*, 797–813.

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

30. Girshick, R., 2015. Fast r-cnn, *Proceedings of the IEEE international conference on computer vision*, 1440–1448.

31. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications, **arXiv preprint arXiv:1704.04861**.

32. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4510–4520.

33. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S., 2017. Feature pyramid networks for object detection, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2117–2125.

34. Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., and Sheikh, Y., 2021. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, **43**.

35. Zhou, X., Karpur, A., Luo, L., and Huang, Q., 2018. Starmap for category-agnostic keypoint and viewpoint estimation, *Proceedings of the European Conference on Computer Vision (ECCV)*, 318–334.

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

37. Smith, B.M., Zhang, L., Brandt, J., Lin, Z., and Yang, J., 2013. Exemplar-based face parsing, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3484–3491.

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

39. Feng, Z.H., Kittler, J., Awais, M., Huber, P., and Wu, X.J., 2018. Wing loss for robust facial landmark localisation with convolutional neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2235–2245.

40. Feng, Y., Wu, F., Shao, X., Wang, Y., and Zhou, X., 2018. Joint 3d face reconstruction and dense alignment with position map regression network, *Proceedings of the European conference on computer vision (ECCV)*, 534–551.

41. Yang, S., Luo, P., Loy, C.C., and Tang, X., 2016. Wider face: A face detection benchmark, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5525–5533.

42. Radosavovic, I., Kosaraju, R.P., Girshick, R., He, K., and Dollár, P., 2020. Designing network design spaces, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10428–10436.

43. Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J., 2018. Path aggregation network for instance segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 8759–8768.

44. Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J., 2021. YOLOX: Exceeding YOLO Series in 2021, **ArXiv**, **abs/2107.08430**.

45. Ge, Z., Liu, S., Li, Z., Yoshie, O., and Sun, J., 2021. Ota: Optimal transport assignment for object detection, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 303–312.

46. Vyas, R., Williams, B.M., Rahmani, H., Boswell-Challand, R., Jiang, Z., Angelov, P., and Black, S., 2022. Ensemble-Based Bounding Box Regression for Enhanced Knuckle Localization, **Sensors**, **22**(4):1569.

47. Zhou, H., Li, Z., Ning, C., and Tang, J., 2017. Cad: Scale invariant framework for real-time object detection, *Proceedings of the IEEE international conference on computer vision workshops*, 760–768.

48. Weng, C.H., Lai, Y.H., and Lai, S.H., 2017. Driver drowsiness detection via a hierarchical temporal deep belief network, *Computer Vision–ACCV 2016 Workshops: ACCV 2016 International Workshops, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part III 13*, 117–133, Springer.

49. Nowara, E.M., Marks, T.K., Mansour, H., and Veeraraghavan, A., 2020. Near-infrared imaging photoplethysmography during driving, **IEEE transactions on intelligent transportation systems**, **23**(4):3589–3600.

50. Kumar, A. and Srikanth, T., 2008. Online personal identification in night using multiple face representations, *2008 19th International Conference on Pattern Recognition*, 1–4, IEEE.

51. Nada, H., Sindagi, V.A., Zhang, H., and Patel, V.M., 2018. Pushing the limits of unconstrained face detection: a challenge dataset and baseline results, *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 1–10, IEEE.

52. Jain, V. and Learned-Miller, E., 2010. FDDB: A Benchmark for Face Detection in Unconstrained Settings, Technical Report UM-CS-2010-009, University of Massachusetts, Amherst.

53. Zhu, X. and Ramanan, D., 2012. Face detection, pose estimation, and landmark localization in the wild, *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2879–2886.

54. Köstinger, M., Wohlhart, P., Roth, P.M., and Bischof, H., 2011. Annotated Facial Landmarks in the Wild: A large-scale, real-world database for facial landmark localization, *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2144–2151.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı, Soyadı	:	Ahmet ÜNLÜHİSARCIKLI
Uyruğu	:	Türkiye Cumhuriyeti
Doğum Tarihi ve Yeri	:	
Telefon	:	+
E-posta	:	
Adres	:	

Derece	**Kurum**	**Mezuniyet Yılı**
Lisans	Bilgisayar Mühendisliği	2020

İŞ DENEYİMİ

Yıl	**Kurum**	**Görev**
2021-Halen		Kurucu Genel Müdür
2022-		Araştırma Görevlisi