



Farklı derin sinir ağı modelleri birleştirilerek yakın kızılötesi görüntülerde yüz tespiti

Face Detection in near infrared images by ensembling different deep neural networks

Ahmet Ünlühisarcıklı^{1,*}, Ahmet Nusret Toprak²

¹ Kayseri Üniversitesi, Yazılım Mühendisliği Bölümü, Kayseri, Türkiye

² Erciyes Üniversitesi, Bilgisayar Mühendisliği Bölümü, Kayseri, Türkiye

Öz

Yüz tespiti, güvenlik, sağlık, endüstri, biyometri gibi alanlarda kritik bir 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 bu çalışmada, yetersiz aydınlatma koşullarında dahi verimli çalışabilen, yakın kızılötesi görüntüler üzerinde yüz tespitine odaklanan bir yaklaşım sunulmaktadır. Bu amaçla öncelikle literatürde iyi bilinen sekiz yüz tespiti derin sinir ağı modeli belirlenmiş ve yakın kızılötesi görüntülerdeki başarımlarını ortaya koymak amacıyla görsel ve sayısal olarak karşılaştırılmıştır. Ayrıca ilgili yöntemlerin çalışma zamanlarının karşılaştırmasına yönelik deneyler CPU ve GPU ortamında gerçekleştirilmiştir. 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 karşılaştırılmıştır. Kombinasyonel Oylama yönteminin, ürettiği %93,6 doğruluk değeriyle literatürdeki sınırlayıcı kutu belirleme yöntemlerinden daha başarılı sonuçlar ortaya koyduğu görülmüştür.

Anahtar kelimeler: Yüz tespiti, Yakın kızılötesi görüntüler, Sınırlayıcı kutu topluluğu birleştirme

1 Giriş

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

Günümüzde yüz tespiti yöntemleri; güvenlik sahasında, sınır kontrolünde, mobil cihazlarda pazarlamada,

Abstract

Face detection plays a crucial role in various areas such as security, healthcare, industry, biometrics, 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, eight 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. Additionally, experiments were carried out in CPU and GPU environments to compare the running times of the face detection methods. 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. Combinational Voting produces more successful results than the other methods in the literature, with an accuracy rate of 93.6%.

Keywords: Face detection, Near-infrared images, Bounding-box ensemble

fotoğrafçılıkta, sağlık sektöründe ve insan-bilgisayar etkileşiminin olduğu daha birçok alanda önemli rol oynamaktadır [1]. Bahsi geçen alanlarda gerçekleştirilen yüz tespiti işlemlerinin, uygun aydınlatma şartlarında yapılabilirdiği 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, girdi görüntülerin yüz tespiti aşaması öncesinde görüntü işleme operatörleri kullanılarak iyileştirmesine yönelik yaklaşımlardır.

Cho vd. çalışmalarında [2], düşük ışık altında yüz tespiti problemine 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 insan vücudu tespit edilmekte, ardından, tespit edilen insanın üst yarısı çıkarılmaktadır. Çıkarılan bölge üzerine, omurga ağ olarak VGG Net-16 modelinin kullanıldığı ve R-CNN yaklaşımının benimsendiği yüz algılayıcısı uygulanarak yüz tespiti yapılmaktadır.

Li vd. [3], gece vakitleri gibi yetersiz aydınlatma koşullarında yapılan yüz tespiti işleminin 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) [4] yöntemi kullanılmıştır. Son olarak ise, PRO-NPD neticesinde elde edilen değerlerin 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. [5] tarafından önerilen Yüksek-Düşük Adaptasyonu (High-Low Adaptation, HLA) başlıklı çalışmada, 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 yaklaşımını sağlamaya çalışmaktadır. Geliştirilen yapı kapsamındaki test ve karşılaştırma süreçlerinde Çift Çekim Yüz Algılayıcısı (Dual Shot Face Detector, DSFD) [6] kullanılmıştır.

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 ortamlarda kullanıcının görüntüsünün gerçek zamanlı olarak alınması gereken durumlarda, gece görüşlü kamera kullanıcıyı rahatsız edici niteliktedir. Diğer taraftan, yetersiz ışık altında da yüz tespiti yapılabilmesini sağlamak amacıyla yakın kızılötesi kameralar kullanılabilir. Bu çalışmada, karanlık gece görüntülerinin iyileştirilmesi ve elde edilen görüntüler üzerinde yüz tespitinin yapılması gaye edinilmiştir.

Dash vd. [7]; farklı kafa pozisyonu, arka plan ve aydınlatma koşulu gibi kızılötesi görüntülerde yüz tespit yöntemlerinin başarımını etkileyebilecek etkenleri araştıran bir çalışma gerçekleştirmiştir. Bu çalışmada, 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 ışık 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 işlemi yapılmıştır. Bu yaklaşımla yapılan yüz tespitinin, çok sayıda insan bulunan kalabalık ortamlarda ve kafa pozisyonlarının farklı olduğu durumlarda başarılı olabileceği öne sürülmüştür.

Gao vd. [8], yetersiz aydınlatma koşullarında görünür spektruma dayalı olarak insan yüzlerinin 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 stratejisi önermiştir. Önerilen stratejideki PDD modeli, ışık yoğunluğunun algılama performansı üzerindeki etkisinin modellenmesiyle elde edilmiştir. Ayrıca, yetersiz aydınlatma koşullarında yüz tespiti performansını daha da artırmak için insan yüzünün merkez noktası tahminine yönelik bir yapay sinir ağı tanıtılmıştır. Merkez noktaya dayalı çalışan bu yöntemin, çapa kutusunun (anchor box) neden olduğu bir dizi sorunu önleyebilecek nitelikte olduğu öne sürülmektedir.

Literatürde renkli ve gri seviye görüntüler üzerinde yüz tespitine yönelik oldukça başarılı sonuçlar veren yöntemler bulunmaktadır. Ancak bu çalışmaların yakın kızılötesi görüntüler üzerinde yüz tespiti başarımının araştırılmasına yönelik bir çalışma bulunmamaktadır. Bu çalışmada, literatürde yaygın olarak kullanılan Dlib [9], FaceBoxes [10], DSFD [6], BlazeFace [11], CenterFace [12], RetinaFace [13], SCRFD [14] ve YuNet [15] yüz tespiti yöntemlerinin yakın kızılötesi görüntüler üzerinde başarımlarının belirlenmesi amacıyla karşılaştırmalı analizleri yapılmıştır.

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) [16], Ağırlıklı Kutular Birleşimi (Weighted Boxes Fusion, WBF) [17] ve Maksimum Olmayan Ağırlıklandırma (Non-Maximum Weighting, NMW) [18] yöntemleri literatürde kullanılan yöntemlerdendir. Bu çalışmada, bahsi geçen sınırlayıcı kutu belirleme yaklaşımlarının kızılötesi görüntülerde yüz tespitinde başarımları karşılaştırılarak performansları değerlendirilmiştir. Ayrıca, yine bu çalışmada; Kombinasyonel Oylama (KO) ve Maksimum

Olmayan Ortalama (Non-Maximum Averaging, NMA) olarak isimlendirilmiş iki yeni sınırlayıcı kutu belirleme yaklaşımı önerilmekte ve literatürdeki yöntemlerle karşılaştırılmaktadır.

Çalışma kapsamında yapılan deneylerde yakın kızılötesi görüntülerden oluşan TUFTS yüz görüntüsü veri seti [19, 20] kullanılmaktadır. Seçilen veri setinde verilen görüntülerde yüz konumları etiketlenmemiş olduğu için veri seti dâhilindeki tüm görüntüler yüz konumlarını gösterecek şekilde etiketlenerek referans değerler (ground truth) oluşturulmuştur. Elde edilen etiketli veriler hem yüz algılayıcılarının hem de sınırlayıcı kutu belirleme yöntemlerinin başarımlarının araştırılmasında kullanılmıştır.

Bu çalışmanın katkılarında birincisi; TUFTS veri seti içerisindeki bütün yakın kızılötesi görüntülerdeki yüz konumlarının sınırlayıcı kutular ile etiketlenmesidir. Çalışmanın katkılarında ikincisi; renkli veya gri seviye görüntülerde başarımları bilinen güncel yüz algılayıcılarının yakın kızılötesi görüntüler üzerindeki performanslarının karşılaştırmasıdır. Üçüncüsü; kullanılan yüz algılayıcılarının yakın kızılötesi görüntülerde bulmuş oldukları yüz konumlarını belirleyen sınırlayıcı kutuların iyileştirilmesi için literatürdeki güncel sınırlayıcı kutu belirleme yöntemlerinin başarımlarının karşılaştırılmasıdır. Yapılan çalışmanın son katkısı ise; Kombinasyonel Oylama (KO) ve Maksimum Olmayan Ortalama (Non-Maximum Averaging, NMA) adları verilen iki yeni sınırlayıcı kutu belirleme yönteminin önerilmesidir.

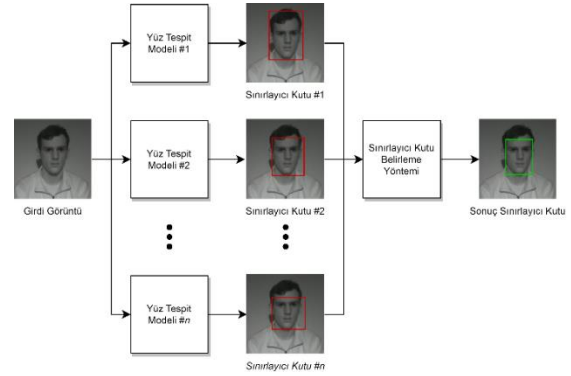
2 Materyal ve metot

Çalışmanın temel amacı olan kızılötesi görüntülerde yüz tespiti başarımlarının araştırılması ve yeni sınırlayıcı kutu belirleme yöntemleri ile bu başarımın artırılmasıdır. Ancak literatürde yüz konumlarının sınırlayıcı kutularla etiklendiği bir veri seti bulunmamaktadır. Bu sebeple; öncelikle TUFTS veri setindeki yakın kızılötesi görüntüler etiketlenmiştir. Kullanılan veri seti, farklı kafa pozisyonlarında alınmış görüntüleri de bünyesinde 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 [21]. Bu sebeple; bu çalışmada derin sinir ağları ile yüz öğrenimini gerçekleştiren ve başarısı kanıtlanmış güncel yüz algılayıcılarından yedi tanesi seçilmiştir. Seçilen modellerden her birinin renkli veya gri seviye görüntüler içeren veri setleri kullanılarak performans analizlerinin yapıldığı çalışmalar mevcuttur [22, 23]. Ancak literatürde bu yöntemlerin yakın kızılötesi görüntüler üzerindeki performanslarının incelendiği bir çalışmaya rastlanamamıştır. Bu sebeple, benzeri çalışmalardan farklı olarak yakın kızılötesi görüntüler üzerinde yüz algılayıcılarının analizinin yapılması ve performanslarını artırıcı nitelikte yeni yaklaşımların önerilmesi amaç

edinilmiştir. Bölüm 2.2'de güncel yüz tespiti yöntemleri tanıtılmaktadır.

Literatürde birden fazla derin sinir ağı modelinin çıktılarını bir araya getirilerek daha başarılı ve kapsayıcı bir sonuç üretmeyi amaçlayan yöntemler bulunmaktadır. Bu yöntemlerden yola çıkarak geliştirilmiş sınırlayıcı kutu belirleme yöntemleri, birden fazla sınırlayıcı kutu tahminini bir araya getirip daha doğru ve daha güvenilir bir sınırlayıcı kutu elde etmeyi amaçlamaktadır. Bu amaç doğrultusunda çalışan dört farklı sınırlayıcı kutu belirleme yöntemi literatürden seçilmiştir. Ayrıca, yakın kızılötesi görüntülerde yüz tespiti yöntemlerinin başarımlarının artırılması amacıyla iki yeni sınırlayıcı kutu belirleme yaklaşımı önerilmektedir. Literatürden seçilen ve önerilen sınırlayıcı kutu belirleme yaklaşımları Bölüm 2.3'te verilmektedir. İlgili yöntemlerin çalışma yapısına Şekil 1'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 yüz konumları çıkarılmaktadır.



Şekil 1. Sınırlayıcı kutu belirleme yöntemi akış diyagramı

2.1 TUFTS yüz görüntüsü veri seti

TUFTS veri seti [19, 20], 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 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, TUFTS veri setinde bulunan bazı yakın kızılötesi görüntü örneklerini göstermektedir. Bu veri setindeki tüm görüntüler bu çalışma kapsamında etiketlenerek, referans sınırlayıcı kutu konumları elde edilmiştir.

İlgili veri setinde yakın kızılötesi görüntüleme için aydınlatma koşulu, 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 oturulmuştur. Kameralar, tripodlara monte edilerek her kameranın yüksekliği, görüntü merkezine karşılık gelecek şekilde manuel olarak ayarlanmıştır. Katılımcıya olan mesafe, görüntü alımı sürecinde sabit tutulmuş ve dağınık ışıklar kullanılarak sabit bir aydınlatma koşulu sağlanması amaçlanmıştır [19].



Şekil 2. TUFTS yüz görüntüsü veri seti yakın kızılötesi örnek görüntüler [20]

TUFTS veri setini etiketleme işlemi, LabelImg aracı [24] kullanılarak yapılmıştır. Bu aşamada; veri seti içerisinde bulunan bütün yakın kızılötesi görüntülerdeki yüz bölgeleri, sınırlayıcı kutu içerisine alınacak şekilde işaretlenmiştir. İşaretlenen her bir görüntü *xml* formatında kaydedilmiştir. Bütün görüntülerin etiketlenmesi sonrasında, alınan veriler tek bir *csv* dosyası içerisinde toplanmıştır. Böylece, aynı formatta etiketlenmiş referans değerleri içeren tek bir dosya elde edilmiştir. Sonraki süreçte, önerilen ve karşılaştırılan tüm yöntemler aynı formatta etiketlenmiş veriler kullanılarak değerlendirilmiştir.

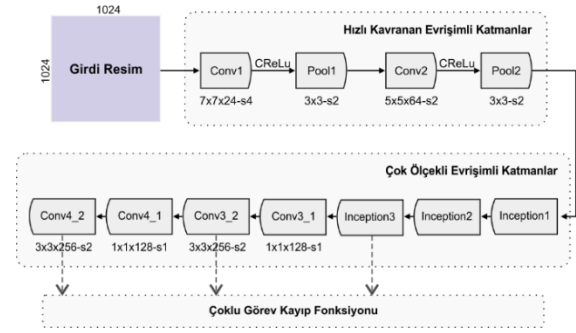
2.2 Yüz tespiti için derin sinir ağı modelleri

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 yöntemleri ve sınırlayıcı kutu belirleme yaklaşımları tanıtılmaktadır. Ayrıca iki yeni sınırlayıcı kutu belirleme yöntemi önerilmektedir.

2.2.1 FaceBoxes

FaceBoxes [10], yalnızca tek bir evrişimli sinir ağı içeren ve uçtan uca eğitilebilen bir yüz algılayıcısıdır. Bu yüz algılayıcısı; Şekil 3'te 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 3. FaceBoxes mimarisi [10]

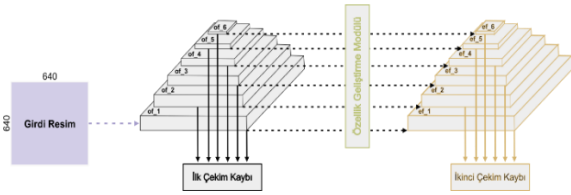
FaceBoxes, görüntü üzerinde farklı ölçeklendirmeler kullanarak farklı boyutlardaki yüz bölgelerini tespit etmeye çalışan kayan pencere (sliding window) yöntemini [25] kullanmaktadır. Bu sayede, farklı ölçeklerdeki çoklu yüzlerin aynı anda tespit edilmesi mümkün hale gelmektedir.

2.2.2 DSFD

Çift Çekim Yüz Algılayıcısı (Dual Shot Face Detector, DSFD) [6]; Tek Çekim Algılayıcısı (Single Shot Detector, SSD) [26] mimarisi üzerine üç yeni değişiklik önererek ortaya çıkan bir yüz algılayıcısıdır. Bu önerilerden ilki; sınıflandırmaların ve regresyonların yürütüldüğü öz nitelik haritalarının hesaplanması üzerine yapılmıştır. Yapılan çalışmada SSD ağından farklılaşan temel unsur, çeşitli derinliklerdeki altı öz nitelik haritasının altı "geliştirilmiş" öz nitelik haritasına dönüştürülmesidir. Önerilerden ikincisi; ağ mimarisinin eğitimi sırasında kayıp fonksiyonu çıktısını en aza indirmeyi hedefleyen bir kayıp fonksiyonunun geliştirilmesidir. Bu yaklaşımda; geleneksel nesne tespit kayıplarından farklı olarak, mimaride yalnızca farklı seviyeler için değil, aynı zamanda çekim olarak tanımlanmış iki farklı yapı için de aşamalı kayıp hesabı uygulanmaktadır. Çift Çekim Yüz Algılayıcısı ismi; bu çift akışlı tasarım sebebiyle verilmiştir. Yüz algılayıcısının 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.

DSFD; 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 4'te yer almaktadır. Ağ yapısı içerisinde, altı farklı evrişim katmanı, ilk çekim algılama katmanları olarak belirlenmiştir. Bu katmanlar, altı adet orijinal öz nitelik (of_1-of_6) haritasını oluşturmaktadır. Daha sonra, önerilmiş yapı içerisinde tek çekim algılayıcısını çift çekim algılayıcısına genişletmek için kullanılan Öz Nitelik Geliştirme Modülü (Feature Enhance Module, FEM) ile bu orijinal öz nitelik haritaları altı gelişmiş öz niteliğe (ef_1-ef_6) aktarılmaktadır. Bu öz nitelik haritaları, orijinal olanlarla aynı boyutlara sahiptir ve ikinci çekim algılama katmanlarını oluşturmak için SSD tarzı bir yaklaşımı benimsemişlerdir.

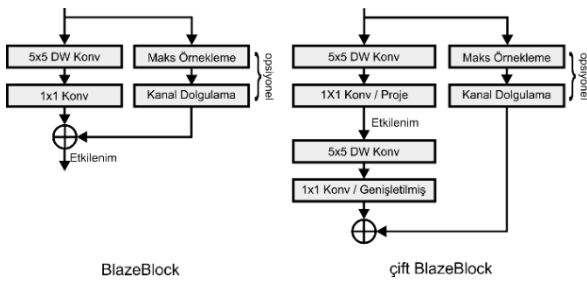


Şekil 4. DSFD mimarisi [6]

2.2.3 BlazeFace

BlazeFace [11], mobil GPU kullanımı için özel olarak tasarlanmış, hafif ve oldukça hızlı çalışan bir yüz algılayıcısıdır. BlazeFace geliştiricileri tarafından, bu modelin mobil cihazlar üzerinde 200-1000 FPS aralığında görüntü işleme kapasitesine sahip olduğu ifade edilmektedir. Bu sayede önerilen model, 2B/3B yüz anahtar noktası veya yüz bölgesi geometri tahmini, yüz özellikleri veya mimik sınıflandırması ve yüz bölgesi bölütlenmesi gibi yüz tespitinin oldukça önemli olduğu problemlerde kullanım alanı bulmaktadır.

Şekil 5'te görüldüğü gibi BlazeFace mimarisinde 5x5 boyutunda derin evrişim kullanılmaktadır. Bu yaklaşımın, tespit sürecini hızlandırmak için modeli daha sığ hale getirdiği düşünülmektedir.



Şekil 5. BlazeFace mimarisi [11]

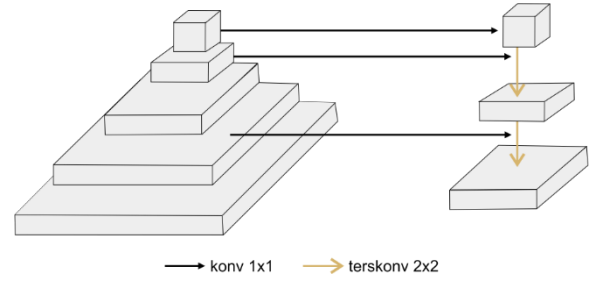
BlazeFace, NMS yöntemine [26, 27, 28] alternatif olarak geliştirilmiş stratejiye ve SSD [26] temel alınarak geliştirilmiş bir modele sahiptir.

2.2.4 CenterFace

CenterFace [12]; gerçek zamanlı hız ve yüksek doğruluk oranıyla, insan yüzü ve yüz karakteristik noktalarının (face landmarks) konumunu aynı anda tahmin etme yetisine sahip bir yüz algılayıcısıdır. Algılayıcının geliştirme sürecinde, özellikle hafif bir ağ yapısı ve daha basit bir yaklaşım

benimsenmiştir. Ağır modellere sahip yüz algılayıcılarının 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 algılayıcısının 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 6'da yer verilmiştir. Mimari içerisinde, omurga ağ olarak MobileNetV2 [29] ve boyun ağları olarak Öz Nitelik Piramidi Ağı (Feature Pyramid Network, FPN) [30] kullanılmıştır. Burada FPN; tek bir ölçek girdisinden tek bir öz nitelik piramidi oluşturmak için, yanıl bağlantılara sahip olan dikey bir mimari yapısına sahiptir. Mimari tasarımı doğrultusunda; yüz algılama ve hizalama için piramitteki yalnızca bir katman kullanılmıştır.

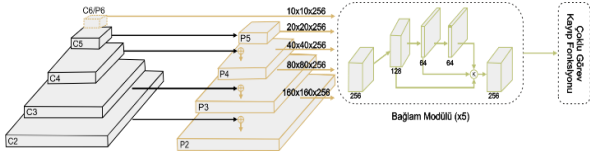


Şekil 6. CenterFace mimarisi [12]

2.2.5 RetinaFace

RetinaFace [13], farklı saha şartlarında verimli bir şekilde yüz bölgesi tespiti gerçekleştirilmesinin zorluğuna değinerek, bu soruna yönelik bir yüz algılayıcısı olarak geliştirilmiştir. Tespit edilen soruna çözüm için geliştirilen yüz algılayıcısı, 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 tabanlı yüz bölgesi tespiti yapabilen güçlü bir yüz algılayıcısı oluşturulmuştur.

RetinaFace mimarisi Şekil 7'de verilmiştir. Mimari incelendiğinde, ilk olarak Öz Nitelik Piramidi Ağına (Feature Pyramid Network, FPN), ardından bağlam modülüne ve son olarak da çoklu görev kayıp fonksiyonuna sahip bir yapı görülmektedir. FPN, 2B görüntü üzerinden beş farklı seviyede öz nitelik çıkarmaktadır. Çıkarmış olduğu öz niteliklerden ilk dördü (C2-C5), önceden eğitilmiş ResNet modeli [31] kullanılarak hesaplanmaktadır. En üstte yer alan en küçük öz nitelik ise (C6), 2 piksellik adımlarla 3x3 boyutunda pencere ile evrişim işlemi uygulanarak çıkarılmaktadır. Elde edilen tüm bu öz nitelikler, daha fazla bağlamsal bilgiye erişebilmek için beş farklı filtreye sahip bağlam modülüne iletilmektedir. Son olarak ise, bu iletiler çoklu görev kayıp fonksiyonuna girdi olarak verilmektedir.



Şekil 7. RetinaFace mimarisi [13]

2.2.6 SCRFD

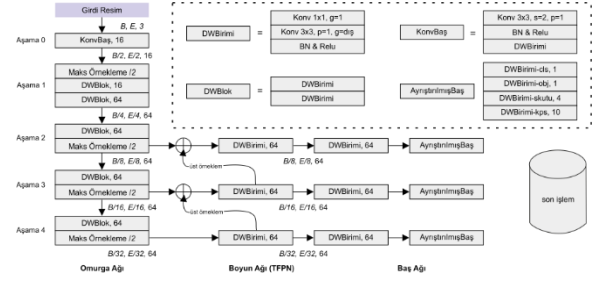
SCRFD [14], düşük çözünürlüklü görüntülerde küçük boyutlu yüzleri algılamanın zorluğuna dikkat çekerek bu soruna odaklanan bir yaklaşım önermiştir. Önerilen yaklaşımda, veri örnekleme (data sampling) ve hesaplama dağılımı (computation distribution) stratejilerinin verimli ve yüksek doğruluk oranlarına sahip yüz tespiti 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 Dağıtım (Sample Redistribution, SR) stratejisidir. İkincisi ise; hesaplamayı modelin omurga ağı, boyun ağı ve baş ağı arasında yeniden paylaşımın Hesaplama Yeniden Dağıtım (Computation Redistribution, CR) stratejisidir. Önerilen SR ve CR stratejileri; birbirini tamamlar niteliktedir. Beraber kullanımlarının başarıyı pozitif yönde doğrudan etkileyeceği düşüncesi ile geliştirilmişlerdir.

Yüz tespiti gerçekleştirme aşamasında bir sınıflandırma ağının 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 [32] kullanılmıştır. Bu arama metodu, ResNet'in [31] omurga ağı olduğu RetinaNet [30] üzerine uygulanmıştır. Boyun ağı olarak Yol Toplama Özelliği Piramit Ağı (Path Aggregation Feature Pyramid Network Network, PAFPN) [33] ve baş ağında 3x3 boyutunda evrişim katmanları kullanılmıştır.

2.2.7 YuNet

Model mimarilerinin ağır olması ve yüksek hesaplama maliyetleri, birçok yüz algılayıcısının 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 algılayıcısı olan YuNet [15] 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'in mimarisi Şekil 8'de verilmiştir. Bir omurga ağı, Küçük Öz Nitelikli Bir Piramit ağı (A Tiny Feature Pyramid Network, TFPN) bulduran boyun ağı ve bir baş ağı olmak üzere ilgili mimari üç bölüme ayrılmıştır.



Şekil 8. YuNet mimarisi [15]

2.3 Sınırlayıcı kutu belirleme yöntemleri

Nesne tespiti yöntemlerinin en sık kullandığı yaklaşımlardan biri bir sınırlayıcı kutu ile tespit edilen nesne konumunun belirlenmesidir. Sınırlayıcı kutular, bir nesnenin koordinatlarını (sol üst ve sağ alt köşelerin piksel konumları) temsil etmektedir. Ancak tek bir sınırlayıcı kutu tahmini, bazen yanıltıcı veya hatalı olabilmektedir. Buradan hareketle geliştirilen sınırlayıcı kutu belirleme yöntemleri, farklı model veya algoritmaların ürettiği sınırlayıcı kutuları kullanarak yeni bir sınırlayıcı kutu konumu belirlemektedir.

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. Bu sayede, farklı yaklaşımların avantajları bir araya getirilerek; daha yüksek doğruluk oranına sahip, daha hassas ve daha güvenilir sonuçlar elde edilebilmektedir.

Bu çalışmada; literatürde mevcut olan Maksimum Olmayan Bastırma, Ağırlıklı Kutular Birleşimi ve Maksimum Olmayan Ağırlıklandırma yöntemleri kullanılarak, girdi olarak verilen sınırlayıcı kutuların iyileştirilmesi ve bu yöntemlerin performanslarının analiz edilmesi amaçlanmaktadır. Buna ilaveten; Kombinasyonel Oylama ve Maksimum Olmayan Ortalama olmak üzere iki adet yeni sınırlayıcı kutu belirleme yöntemi önerilmektedir.

2.3.1 Maksimum olmayan bastırma

Maksimum Olmayan Bastırma (Non-Maximum Suppression, NMS) [16], bir yüz/nesne algılayıcısı tarafından oluşturulmuş ve birbiri ile örtüşen sınırlayıcı kutuları indirgemek için kullanılan tekniktir. Yüz algılama, nesne takibi, görüntü ve videolarda nesne algılama gibi birçok alandaki farklı görevlerde yaygın olarak kullanılmaktadır.

NMS yönteminin arkasındaki temel fikir, en yüksek güven skoruna sahip sınırlayıcı kutuyu seçmektir. Bu seçim süreci, en yüksek skora sahip olan sınırlayıcı kutu ile kullanıcı tarafından belirlenen eşik değerinden daha yüksek oranda örtüşen diğer sınırlayıcı kutuları elemek 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; belirli nesnelere doğru olarak tespit edilmesini sağladığı gibi yanlış nesne tespitine de yol açmaktadır. NMS, yalnızca tespit edilen nesnelere karşılık gelen en ilgili sınırlayıcı

kutuları seçerek bu yanlış pozitifleri ortadan kaldırmayı amaçlamaktadır. Yanlış veya gereksiz tespiti ortadan kaldıran bu yöntem, nesne tespit algoritmalarının hesaplama karmaşıklığını azaltmaya da yardımcı olmaktadır.

2.3.2 Ağırlıklı kutular birleşimi

Ağırlıklı Kutular Birleşimi (Weighted Boxes Fusion, WBF) [17], 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 skorlarına dayalı olarak sınırlayıcı kutu koordinatlarının ilgili modelin ağırlığı temel alınarak birleştirilmesini içermektedir. Buna göre her yöntemin ürettiği sınırlayıcı kutu için, ilgili yöntemin güven skoru ile orantılı olarak bir ağırlık değeri belirlenmektedir. Sonuç sınırlayıcı kutu konumu, belirlenen ağırlık değerleri göz önünde bulundurularak, sınırlayıcı kutuların koordinat değerlerinin ağırlıklı ortalaması alınarak hesaplanmaktadır. Daha yüksek güven skoruna sahip sınırlayıcı kutuya daha yüksek ağırlıklar atayarak; güven skoru yüksek tahminlerden daha fazla etkilenen, düşük güven skorlu veya hatalı tahminlerden ise daha az etkilenen sonuç birleştirilmiş sınırlayıcı kutu üretilmektedir.

WBF yönteminin amacı, farklı modellerin güçlü yanlarını birleştirerek daha güçlü bir sonuç elde etmektir. Ayrıca, yanlış tespit noktalarının veya aykırı değerlerin etkisini azaltılarak nesne algılama modellerinin performansı artırılmaktadır.

2.3.3 Maksimum olmayan ağırlıklandırma

Maksimum Olmayan Ağırlıklandırma (Non-Maximum Weighting, NMW) [18], 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, Maksimum Olmayan Bastırma (NMS) tekniğinden esinlenilerek geliştirilmiştir. NMS, gereksiz sınırlayıcı kutuları elemek ve yalnızca en güvenilir ve örtüşmeyen tahminleri tutmak için kullanılmaktadır. NMS yaklaşımında, en yüksek güven skoruna sahip sınırlayıcı kutu ile diğer her bir sınırlayıcı kutu IoU metriği kullanılarak karşılaştırılmaktadır. Ardından, belirlenen eşik değerin üzerinde örtüşmeye sahip sınırlayıcı kutu elimine edilmektedir. Bu durum, aynı nesne üzerinde yinelenen veya oldukça benzer konumlandırılan sınırlayıcı kutuları ortadan kaldırmaya yöneliktir. Bu yaklaşım, belirlenecek son sınırlayıcı kutunun iyileştirilmesini sağlamaktadır. NMW yönteminde ise; gerçekleştirilen eleme işlemi sırasında ek ağırlıklandırma faktörleri dâhil edilerek elde edilecek çıktı performansının yükseltilmesi hedeflenmektedir. Bu faktörler; hangi sınırlayıcı kutuların saklanacağını veya eleneceğini belirlerken, güven skorları, nesne boyutları veya diğer olası niteliklerin dikkate alınmasını içermektedir.

Literatürde yapılan çalışmalar [34], 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 kullanılabilir. Bu yöntem ayrıca, nesne takibi, nesne tanıma ve diğer benzeri uygulamalarda da kullanılabilir.

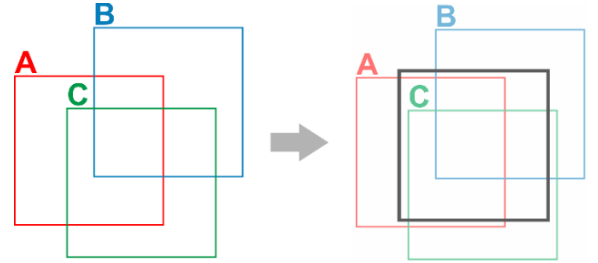
2.3.4 Ortalama alma

Sınırlayıcı kutu belirleme amacıyla kullanılacak en basit yaklaşım olan Ortalama Alma (OA), sınırlayıcı kutu koordinatlarının ortalamasının alınmasını benimseyen sınırlayıcı kutu belirleme yöntemidir. İlgili yöntem; bütün sınırlayıcı kutulara eşit mesafede yaklaşmakta olup herhangi bir ağırlıklandırma durumu söz konusu değildir. Bu yöntem; girdi olarak verilen her bir sınırlayıcı kutunun köşe koordinatlarının ayrı ayrı ortalamasını almaktadır. Şekil 9'da 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:

$$B_{sonuç,j} = \frac{1}{n} \sum_{i=1}^n B_{i,j}, \quad (1)$$

burada n , girdi olarak verilen sınırlayıcı kutu sayısını, $B_{i,j}$ i . sınırlayıcı kutunun j . köşesinin koordinatlarını ifade etmektedir.

Birbirine yakın konumlarda üretilmiş sınırlayıcı kutular üzerinde etkili sonuçlar veren yöntem; konumlandırılmaları birbirinden uzak veya bağımsız sınırlayıcı kutular üzerinde etkili çalışmamaktadır. Bu sebeple; yöntemin kullanımı öncesinde, Jaccard (IoU) metriği ile hedef sınırlayıcı kutuların ön analizi sonrası, keşif oranı 0,8 ve üzeri olan durumlarda kullanımı tavsiye edilmektedir.



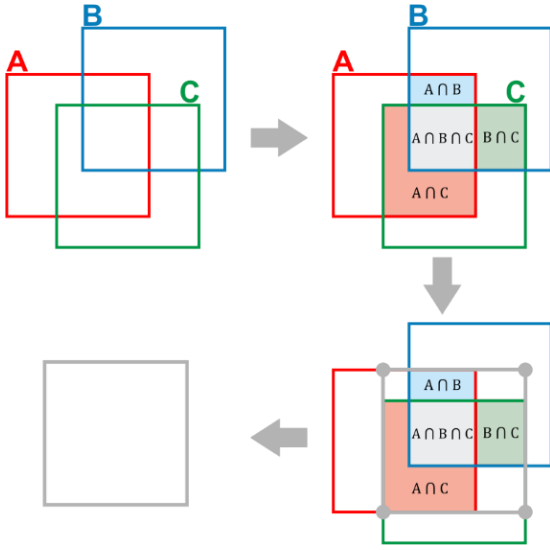
Şekil 9. Ortalama alma yöntemi çalışma şeması

2.3.5 Kombinasyonel oylama

Önerilen ilk sınırlayıcı kutu belirleme yöntemi olan Kombinasyonel Oylama (KO), sınırlayıcı kutu topluluğu kombinasyonlarının oylama yöntemi temel alınarak değerlendirilmesine odaklanmaktadır. Bu yöntem, farklı modellerinin tahminlerini bir araya getirerek en çok tercih edilen tahmini seçmeye dayanmaktadır.

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 10'da 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

bölgesinin tespiti için ya kutulardan iki tanesinin ya da üç tanesinin seçim alanlarını barındıran bölgeler aranmaktadır. Bu arama sürecinde, Şekil 10'da 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 10. 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:

$$IoU(B_i, B_j) = \frac{B_i \cap B_j}{B_i \cup B_j} \quad (2)$$

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, liste içerisindeki diğer sınırlayıcı kutular ile IoU metriği ile analiz edilmektedir:

$$IoU(B_i, B_{maks}) = \frac{B_i \cap B_{maks}}{B_i \cup B_{maks}} \quad (3)$$

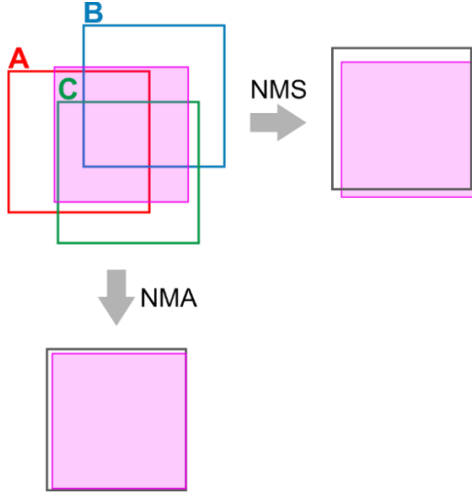
Analiz neticesinde bulunan değer, belirlenen eşik değerden küçük ise listeden kaldırılmaktadır. Bu çalışmada eşik değer 0,8 olarak belirlenmiştir. Bu sayede, farklı kombinasyonlar ile daha yüksek doğruluk oranına sahip sonuç sınırlayıcı kutunun üretildiği oylama yöntemine katkı sağlayan karar verici 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 13 '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 yanlıcı 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.

2.3.6 Maksimum olmayan ortalama

Geliştirilen ikinci yöntem olan Maksimum Olmayan Ortalama (Non-Maximum Averaging, NMA) yöntemi, Maksimum Olmayan Bastırma (NMS) ve Maksimum Olmayan Ağırlıklandırma (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, son olarak en yüksek güven skoruna sahip sınırlayıcı kutu sonuç olarak çıkarılmaktadır. NMW ise; NMS yönteminin gerçekleştirdiği bastırma işlemi sırasında ek ağırlıklandırma faktörlerini kullanarak sonuç sınırlayıcı kutunun doğruluk oranını artırmak üzere geliştirilmiştir. Bu faktörler; hangi sınırlayıcı kutuların saklanacağını veya bastırılacağını belirlemede aktif rol oynamaktadır. Bu sayede, NMS yönteminin güven skorları üzerine türetilmiş olduğu yaklaşıma ağırlık faktörü eklenerek yeni bir yöntem öne sürülmüştür. Buradaki yaklaşımdan esinlenilerek maksimum olmayan ortalama yöntemi önerilmektedir. Önerilen yöntem, NMS yönteminin en yüksek güven skorlu sınırlayıcı kutuyu seçme stratejisini ortalama alma tekniğiyle ileriye taşımaktadır. İlâveten, 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 11'deki şemada yer almaktadır.



Şekil 11. 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:

$$I_i = IoU(B_i, B_{maks}), \quad (4)$$

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 ($B_{sonuç}$) uygulanan bu ortalama yaklaşımı ile elde edilmektedir:

$$B_{sonuç} = \frac{\sum_{i=1}^n B_i I_i}{\sum_{i=1}^n I_i} \quad (5)$$

3 Bulgular ve tartışma

3.1 Deney altyapısı

Bu çalışma kapsamında yapılan tüm deneyler, Intel Core i9 10850K 3,60 GHz 20MB önbelleğe sahip işlemciye, RTX 3060 12GB 192 bit NVIDIA ekran kartına 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 çalışmada, Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Skor (F1 Score) ve Jaccard (IoU) değerlendirme metrikleri kullanılmış olup her biri aşağıda açıklanmıştır:

1. Doğruluk: Toplam veri sayısının ne kadar doğru tespit edilebildiğini bulmaya odaklı bir metriktir. Doğru yapılan tahmin sayısının toplam tahmin (gerçek pozitif (TP), gerçek negatif (TN), yanlış

pozitif (FP), yanlış negatif (FN)) sayısına oranlanması ile hesaplanmaktadır:

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

2. Kesinlik: Gerçek pozitifler (TP) ile tüm pozitiflerin oranıdır:

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (7)$$

3. Duyarlılık: Gerçek pozitiflerin tüm doğru tespitlere oranını göstermektedir:

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (8)$$

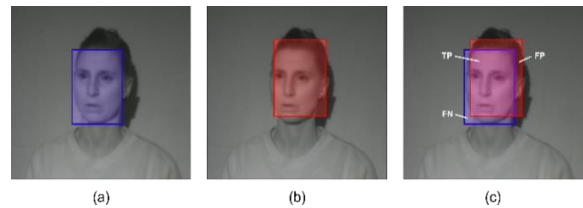
4. F1 Skor: Kesinlik ve Duyarlılık performans değerlendirme metriklerinin harmonik ortalamasının hesaplanması ile elde edilmektedir:

$$\text{F1 Skor} = \frac{2TP}{2TP + FP + FN} \quad (9)$$

5. Jaccard: İki kümenin benzerliğini ölçmek için kullanılan bir metriktir. Bu metrik, iki nesnenin (N_1, N_2) kesişim kümesinin, birleşim kümesine oranını hesaplamaktadır:

$$\text{Jaccard}(N_1, N_2) = \frac{N_1 \cap N_2}{N_1 \cup N_2} \quad (10)$$

Kullanılan performans değerlendirme metrikleri, referans değerler ile modeller tarafından tahmin edilen değerlerin karşılaştırılması suretiyle çalışmaktadır. Bu çerçevede, Şekil 12 (a), referans değer ile oluşturulmuş gerçek yüz bölgesini, Şekil 12 (b), yüz tespit yöntemi tarafından tahmin edilen yüz bölgesini, Şekil 12 (c) ise (a) ve (b)'nin beraber değerlendirilmesi ile elde edilen TP, FN ve FP alanları ifade etmektedir.



Şekil 12. (a) gerçek yüz bölgesi, (b) tahmin edilen yüz bölgesi, (c) yapılan tahmin için TP, FN ve FP alanlar

3.3 Yüz Tespiti derin sinir ağı modellerinin yakın kızıl ötesi görüntüler üzerinde karşılaştırılması

Derin sinir ağı modelleri, büyük miktarda veriyi kullanarak karmaşık özellikleri öğrenme kabiliyetine sahiptirler. Bu kabiliyetlerinden dolayı, görüntülerdeki yüzleri tespit etmek ve tanımak için yaygın olarak kullanılırlar. Derin sinir ağı modelleri ile geliştirilmiş yüz algılayıcıları 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,

bölümde, literatürde eksikliği tespit edilmiş kızılötesi görüntüler üzerinde yüz tespiti senaryosunda, farklı yüz algılayıcılarının 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

(Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Skor ve Jaccard performans değerlendirme metrikleri kullanılarak değerlendirilmiştir. Düz kafa pozisyonundaki resimlerden elde edilen sonuç değerleri **Tablo 1**'de, yan kafa pozisyonundaki sonuç değerleri **Tablo 2**'de ve tamamının beraber ele alındığı değerler ise **Tablo 3**'te yer almaktadır.

Tablo 1. Yüz algılayıcılarının düz kafa pozisyonunda karşılaştırılması

Yüz Algılayıcı	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 2. Yüz algılayıcılarının yan kafa pozisyonunda karşılaştırılması

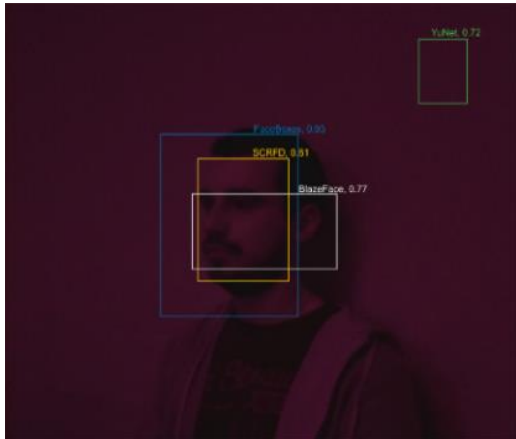
Yüz Algılayıcı	Doğruluk	Kesinlik	Duyarlılık	F1 Skor	Jaccard
BlazeFace	0.5913	0.6126	0.8435	0.6979	0.5550
CenterFace	0.8495	0.8205	0.9366	0.8658	0.7690
Dlib (HOG)	0.5417	0.6038	0.5984	0.5969	0.4858
DSFD	0.9162	0.9183	0.9313	0.9224	0.8589
FaceBoxes	0.8206	0.7861	0.9244	0.8417	0.7436
RetinaFace	0.9202	0.9237	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. Yüz algılayıcılarının bütün görüntüler üzerinde karşılaştırılması

Yüz Algılayıcı	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

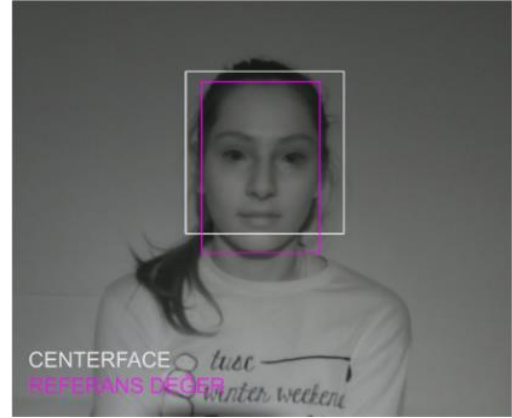
Tablolar incelendiğinde, yan kafa pozisyonlarında gerçekleştirilen yüz tespiti işlemlerindeki başarının düz kafa pozisyonuna göre daha düşük olduğu görülmektedir. Bu durum, yüz algılayıcılarının bütün görüntüler üzerindeki başarısını da doğrudan etkileyerek farklı kafa pozisyonlarına karşı tutarlılıklarını ortaya koymaktadır. Örneğin; MobileNet kullanılarak eğitimi gerçekleştirilmiş RetinaFace yüz algılayıcısı; düz kafa pozisyonlarında %89 doğruluk oranına sahipken, yan kafa pozisyonlarında bu oran %64'e düşmektedir. Aradaki bu büyük fark, ilgili yüz algılayıcısının yakın kızılötesi görüntülerde tutarlı çalışmadığını göstermektedir. Diğer taraftan, seçilen modeller içerisinde en yüksek doğruluk oranlarına sahip olan SCRFD yüz algılayıcısı, düz kafa pozisyonunda %93, yan kafa pozisyonunda ise %92 doğruluk oranı ile yüz tespiti yapmıştır. İlgili yüz algılayıcı, 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 algılayıcı olmuştur. Seçilen modeller arasındaki en güncel yüz algılayıcısı olan YuNet ise, en düşük doğruluk oranlarına sahip olmuş ve yakın kızılötesi görüntüler üzerinde yeterli tespit başarısı gösterememiştir.

Literatürde yer alan her yüz algılayıcısının başarı değerlendirme kriterleri ve kullandığı referans değerler birbirinden farklıdır. Bu durum, yüz algılayıcılarının ürettiği güven skorlarının yer aldığı Şekil 13'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 algılayıcılarının ürettiği 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.



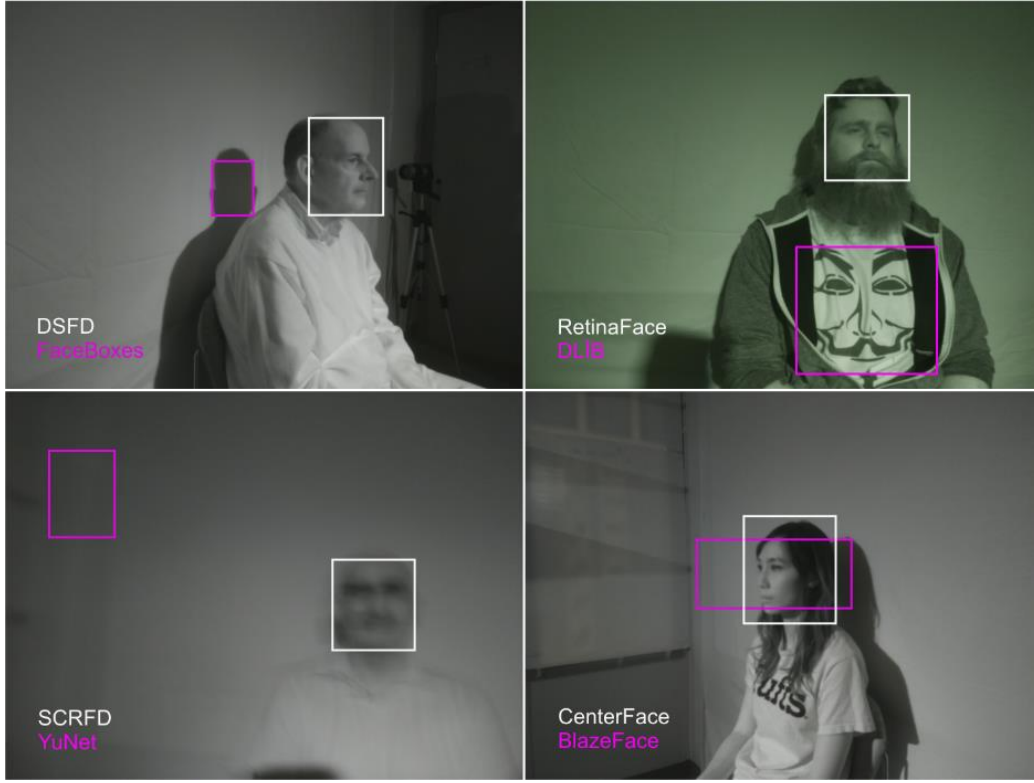
Şekil 13. Yüz algılayıcıları tarafından bulunan yüz bölgeleri ve ürettikleri güven skorları

Yüz tespiti sonrası örnek bir sonuç görüntüsüne Şekil 14'te yer verilmiştir. Burada, etiketleme işlemi sonrasında üretilen referans değer ve CenterFace yüz algılayıcısı tarafından tespit edilen yüz bölgeleri görülmektedir. İlgili şekilde görüleceği üzere, referans değer tam yüz bölgesini çerçevelemektedirken, CenterFace yüz alanının bazı bölgelerini tam alamamakta ve enlemesine daha geniş olarak bulmaktadır. Bu durum tespit edilemeyen noktalar üzerinde yapılacak çalışmaları başarısız, diğer alanlarda gerçekleştirilecek işlemleri ise daha maliyetli yapacak niteliktedir. Bu sebeple, yüz bölgesinin net olarak tayin edilmesi büyük önem taşımaktadır. Bu doğrultuda, renkli ve gri seviye görüntülerde başarılı sonuçlar veren yüz algılayıcılarının 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 sergilenmektedir.



Şekil 14. Referans değer ve CenterFace tarafından çıkarılan yüz bölgeleri

Bazı yüz algılayıcıları tarafından tespit edilen yüz bölgelerinin karşılaştırmalı ele alındığı görsellere Şekil 15'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. Tablo 1-3'te verilen tespit başarı oranlarını direkt etkileyen bu durum, yüz tespiti sonrası gerçekleştirilebilecek yüz tanıma, kafa pozisyon tahmini, duyu 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 algılayıcılarının 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 yüz algılayıcısının en yüksek doğruluk oranlarına, YuNet yüz algılayıcısının ise en düşük doğruluk oranlarına sahip olması durumu anlaşılmaktadır.



Şekil 15. Farklı yüz algılayıcıları tarafından tespit edilen yüz bölgeleri

3.4 Sınırlayıcı kutu belirleme yöntemlerinin yakın kızıl ötesi görüntüler üzerinde 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 kullanılabilir 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 yapılan analizlerde Bölüm 3.2.'de yapılan deneylerde en iyi sonucu veren yüz tespiti derin sinir ağı modelleri kullanılmıştır. Bu bağlamda en iyi sonucu veren modeller beşli, dördü, 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. Gerçekleştirilen çalışmaya ilişkin sonuç değerler Tablo 4-7'de yer almaktadır.

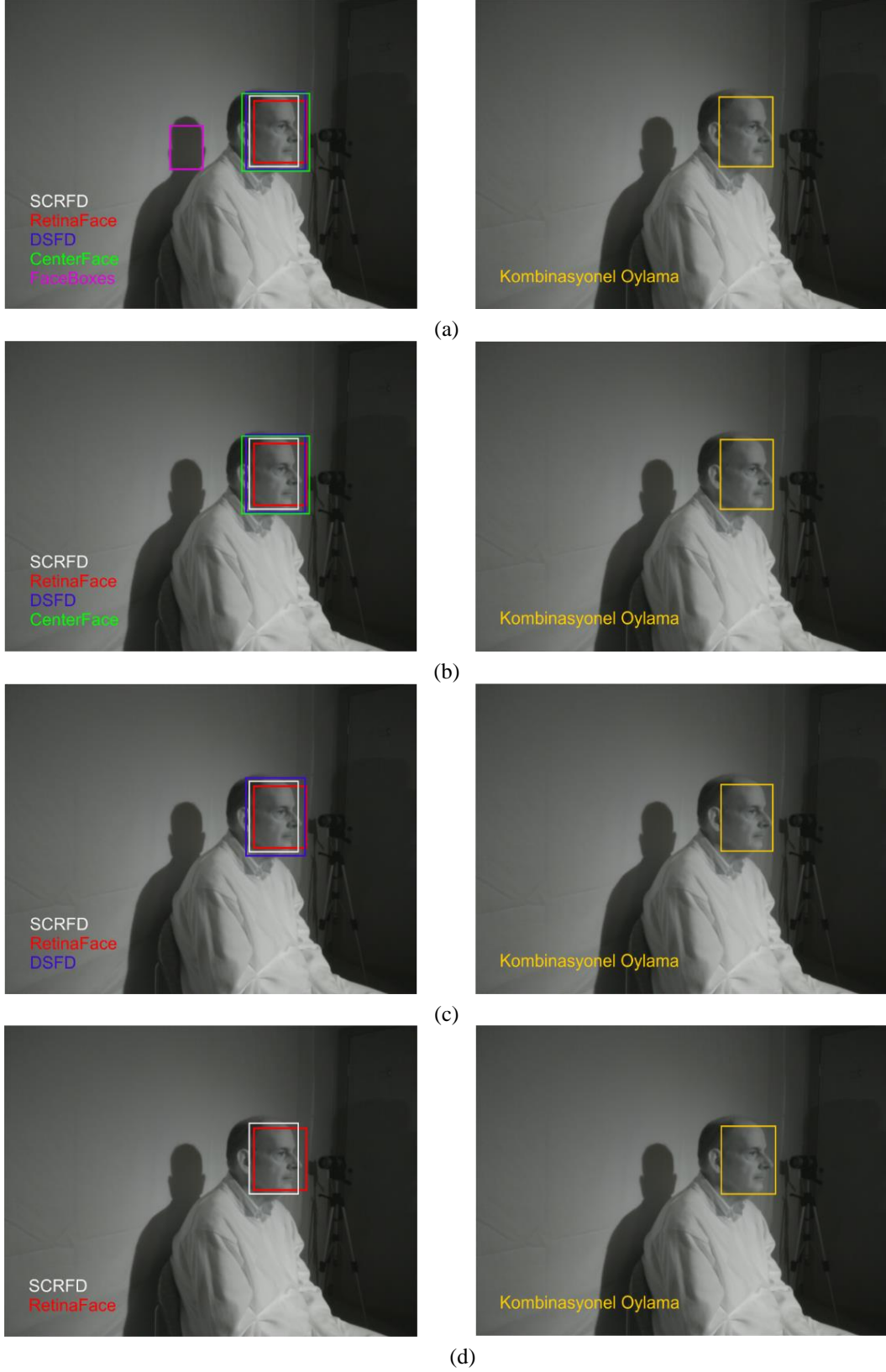
Sayısal sonuçlar incelendiğinde; en iyi beş modelin (Tablo 4) kullanıldığı senaryoda; önerilen KO yöntemi Kesinlik haricindeki bütün performans değerlendirme

metriklerinde en başarılı sonucu üretmiştir. En iyi dört modelin (Tablo 5) kullanıldığı senaryoda, Doğruluk ve Duyarlılık metriklerinde, en iyi üç modelin (Tablo 6) kullanımında Doğruluk ve Jaccard metriklerinde, en iyi iki modelin (Tablo 7) 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 ve Tablo 4 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. İlgili yöntemin farklı senaryolarda ürettiği sonuç sınırlayıcı kutular Şekil 16'da yer alan görsellerde her bir girdi durumu için ayrı ayrı verilmiştir. Örneğin Şekil 16 (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 bölgesini tam olarak alınmamaktadır. Diğer taraftan, KO yönteminin beş farklı modelin ürettiği sınırlayıcı kutuları kullanarak tüm modellerden daha başarılı bir sınırlayıcı kutu ürettiği görülmektedir.

Karşılaştırmalar neticesinde; sınırlayıcı kutu belirleme yöntemlerinin birçoğunun mevcut yüz algılayıcılarının başarısından daha yüksek sonuçlar ürettiği görülmektedir.

Bu durum sınırlayıcı kutu belirleme yöntemlerinin mevcut yüz veya nesne algılayıcılarının tespit doğruluğunu artırıcı nitelikte olduğunu kanıtlamaktadır.



Şekil 16. En iyi sonucu veren beş (a), dört (b), üç (c) ve iki (d) modelin kullanımı ile kombinasyonel oylama yönteminden yeni sınırlayıcı kutu üretilmesi

Tablo 4. En iyi sonucu veren beş modelin (SCRFD, RetinaFace, DSFD, CenterFace, FaceBoxes) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınırlayıcı 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 5. En iyi sonucu veren dört modelin (SCRFD, RetinaFace, DSFD, CenterFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınırlayıcı 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	0.9357	0.8771
KO	0.9337	0.9066	0.9689	0.9348	0.8749

Tablo 6. En iyi sonucu veren üç modelin (SCRFD, RetinaFace, DSFD) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınırlayıcı 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	0.9310	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 7. En iyi sonucu veren iki modelin (SCRFD, RetinaFace) girdi olarak verildiği durumda sınırlayıcı kutu belirleme yöntemlerinin karşılaştırılması

Sınırlayıcı 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

3.5 Yüz tespiti derin sinir ağı modellerinin işlem zamanlarının karşılaştırılması

Modellerin işlem zamanlarının karşılaştırılması, farklı modellerin aynı işlemi gerçekleştirmek için tükettiği süreleri ve performanslarını değerlendirmek amacıyla yapılmıştır. Yapılan çalışma hem CPU hem de GPU üzerinde yürütülerek, Tablo 8'de 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.

Gerçekleştirilen değerlendirme sürecinde; işlem zamanı karşılaştırması ile, ilgili modellerin kullandığı omurga ağı, sahip olduğu parametre sayısı ve kayan nokta operasyonu (Floating Point Operations, FLOP) miktarları da göz önüne alınmıştır. Bahsedilen çalışmaya ilişkin karşılaştırma verileri Tablo 9'da yer almaktadır. Sonuç olarak; Tablo 8 ve Tablo 9 beraber incelendiğinde, modellerin sahip oldukları parametre sayıları ve FLOP miktarları ile işlem zamanları arasında ilişki olduğu görülmektedir.

Tablo 8. İşlem zamanı karşılaştırma tablosu

Yüz Algılayıcı	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	-

Tablo 9. Modellere ait omurga ağı, parametre ve FLOP karşılaştırma tablosu

Yüz Algılayıcı	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	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

4 Sonuçlar

Yetersiz ışık altında verimli bir 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 algılayıcısının 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 algılayıcıların yakın kızılötesi görüntülerde istikrarlı çalışma durumları ve tutarlılıkları ortaya çıkarılmıştır. Bunlara ilaveten, ilgili yüz algılayıcıların 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.

Yüz algılayıcıların 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 algılayıcılarının daha yüksek başarıya sahip olması ile örtüşmektedir. Bütün kafa pozisyonlarında %93 ile en yüksek doğrulukta tespit yapan SCRFD, yakın kızılötesi görüntüler üzerinde en başarılı yüz tespiti işlemi gerçekleştiren model olmuştur. Seçilen modeller arasında en güncel yüz algılayıcısı olan YuNet ise, bütün kafa pozisyonlarında %42 doğruluk oranı ile en düşük başarıyı gösteren yüz tespiti yöntemi konumundadır. Diğer taraftan, Retina_MobileNet, BlazeFace ve Dlib düz kafa pozisyonunda sırasıyla; %89, %75 ve %79 doğruluk oranlarıyla daha başarılı sonuçlar üretirken, yan kafa pozisyonunda elde ettikleri %64, %59 ve %54 doğruluk değerleri ile yeterli başarıyı sağlayamamıştır. Bu yüz algılayıcıları 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 algılayıcısı 0.011 saniye ile BlazeFace, en yavaş çalışan ise 3 saniye ile Dlib olmuştur. GPU üzerinde; en hızlı çalışan model olan FaceBoxes 0,004 saniyede yüz tespiti yaparken, en yavaş çalışan model olan RetinaFace 0,8 saniyede yüz tespiti yapmıştır. Burada, RetinaFace yüz algılayıcısının 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 algılayıcılarının 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, sırasıyla ortalama doğruluk değerleri olan %92,6 ve %92,5 oranlarının altında kalmıştır. Bu durum anılan yöntemlerin bazı senaryolarda basit ortalama alma yaklaşımının dahi altında kalarak beklenen 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 algılayıcıları tarafından üretilen tespit noktalarının başarısını artıracak niteliktedir. Ayrıca, sunulan yöntemlerden Kombinasyonel Oylama yönteminin, mevcut sınırlayıcı kutu belirleme yöntemlerinden daha yüksek başarıya sahip olduğu görülmüştür.

Teşekkür

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

Çıkar çatışması

Gerçekleştirilen çalışmalar kapsamında herhangi bir çıkar çatışması bulunmamaktadır.

Benzerlik oranı (iThenticate): %5

Kaynaklar

- [1] Kumar, A., Kaur, A., & Kumar, M. Face detection techniques: a review. *Artificial Intelligence Review*, 52, 927-948. 2019.
- [2] Cho, S. W., Baek, N. R., Kim, M. C., Koo, J. H., Kim, J. H., & Park, K. R. Face detection in nighttime images using visible-light camera sensors with two-step faster region-based convolutional neural network. *Sensors (Switzerland)*, 18(9). 2018. <https://doi.org/10.3390/s18092995>
- [3] Li, J., Zhang, D., Zhang, K., Hu, K., & Yang, L. Real-time face detection during the night. 2017 4th International Conference on Systems and Informatics, ICSAI 2017, 2018-January. <https://doi.org/10.1109/ICSAI.2017.8248358>
- [4] S Liao, A K Jain and S Z. Li, "A fast and accurate unconstrained face detector[J]", *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 38, no. 2, pp. 211-223, 2016.
- [5] Wang, W., Wang, X., Yang, W., & Liu, J. Unsupervised face detection in the dark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1), 1250-1266. 2023. <https://doi.org/10.1109/TPAMI.2023.3152562>
- [6] Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J., & Huang, F. DSFD: Dual shot face detector. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019-June, 5055-5064. <https://doi.org/10.1109/CVPR.2019.00520>
- [7] Dash, P., Kisku, D. R., Gupta, P., & Sing, J. K. Fast face detection using a unified architecture for unconstrained and infrared face images. *Cognitive Systems Research*, 74, 18-38. 2022.
- [8] Gao, Z., Zhang, S., Fang, H., Li, L., & Huang, L. Multi-modal image fusion based improved face detection algorithm in poor lighting conditions. *ACM International Conference Proceeding Series*. 2021. <https://doi.org/10.1145/3483207.3483213>
- [9] King, D. E. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10, 2009.
- [10] Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., & Li, S. Z. FaceBoxes: A CPU real-time face detector with high accuracy. *IEEE International Joint Conference on Biometrics, IJCB 2017, 2018-January*, 1-9. 2018. <https://doi.org/10.1109/BTAS.2017.8272675>
- [11] Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K., & Grundmann, M. BlazeFace: Sub-millisecond neural face detection on mobile GPUs. 2019. <http://arxiv.org/abs/1907.05047>
- [12] Xu, Y., Yan, W., Yang, G., Luo, J., Li, T., & He, J. CenterFace: Joint face detection and alignment using face as point. *Scientific Programming*, 2020, 1-8. <https://doi.org/10.1155/2020/7845384>
- [13] Deng, J., Guo, J., Ververas, E., Kotsia, I., & Zafeiriou, S. (2020). RetinaFace: Single-shot multi-level face localisation in the wild. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5202-5211. 2020. <https://doi.org/10.1109/CVPR42600.2020.00525>
- [14] Guo, J., Deng, J., Lattas, A., & Zafeiriou, S. Sample and computation redistribution for efficient face detection. *ICLR 2022 - 10th International Conference on Learning Representations*. 2022.
- [15] Wu, W., Peng, H., & Yu, S. YuNet: A tiny millisecond-level face detector. *Machine Intelligence Research*. 2023. <https://doi.org/10.1007/s11633-023-1423-y>
- [16] Neubeck, A., & van Gool, L. Efficient non-maximum suppression. *Proceedings - International Conference on Pattern Recognition*, 3. 2006. <https://doi.org/10.1109/ICPR.2006.479>
- [17] Solovyev, R., Wang, W., & Gabruseva, T. Weighted boxes fusion: Ensembling boxes from different object detection models. 2019. <https://doi.org/10.1016/j.imavis.2021.104117>
- [18] C. Ning, H. Zhou, Y. Song, and J. Tang. Inception single shot multibox detector for object detection. In *ICME*, 2017.
- [19] Panetta, Karen, Qianwen Wan, Sos Ağaian, Srijith Rajeev, Shreyas Kamath, Rahul Rajendran, Shishir Rao et al. "A comprehensive database for benchmarking

- imaging systems." IEEE Transactions on Pattern Analysis and Machine Intelligence. 2018.
- [20] TUFTS Face DB, TUFTS yüz görüntüsü veri seti depolama alanı, <http://tdface.ece.tufts.edu/>, son erişim 23 Kasım 2023.
- [21] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. Object detection in 20 years: A survey. Proceedings of the IEEE. 2023.
- [22] Mamieva, D., Abdusalomov, A. B., Mukhiddinov, M., & Whangbo, T. K. Improved face detection method via learning small faces on hard images based on a deep learning approach. Sensors, 23(1), 502. 2023.
- [23] Adouani, A., Henia, W. M. ben, & Lachiri, Z. A comparison of face detection methods using spontaneous videos. Multimedia Tools and Applications, 81(16), 23163–23191. 2022. <https://doi.org/10.1007/s11042-022-12781-8>
- [24] LabelImg, Tzutalin LabelImg Git code (2015). <https://github.com/tzutalin/labelImg>, son erişim 24 Kasım 2023.
- [25] R. Vaillant, C. Monrocq, and Y. Le Cun. Original approach for the localisation of objects in images. IEE Proceedings Vision, Image and Signal Processing, 1994.
- [26] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. SSD: Single shot multibox detector. Lecture Notes in Computer Science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 9905 LNCS. 2016. https://doi.org/10.1007/978-3-319-46448-0_2
- [27] Girshick, R. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448). 2015.
- [28] Ren, S., He, K., Girshick, R., & Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 2015-January.
- [29] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2018. <https://doi.org/10.1109/CVPR.2018.00474>
- [30] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. Feature pyramid networks for object detection. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January. <https://doi.org/10.1109/CVPR.2017.106>
- [31] He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December. <https://doi.org/10.1109/CVPR.2016.90>
- [32] Radosavovic, I., Kosaraju, R. P., Girshick, R., He, K., & Dollár, P. Designing network design spaces. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.2020, <https://doi.org/10.1109/CVPR42600.2020.01044>
- [33] Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. Path Aggregation Network for Instance Segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.2018. <https://doi.org/10.1109/CVPR.2018.00913>
- [34] Zhou, H., Li, Z., Ning, C., & Tang, J. CAD: Scale invariant framework for real-time object detection. proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017, 2018-January. <https://doi.org/10.1109/ICCVW.2017.95>

