

Yüz Görüntülerine Morflemeye Dayalı Maske Giydirmeye ve Maskeli Yüz Tanıma

Asuman GÜNAY YILMAZ¹, Eyüp GEDİKLİ¹, Omar ALHORİ¹

¹ Karadeniz Teknik Üniversitesi, Of Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Trabzon, Türkiye

(Alınış / Received: 20.01.2022, Kabul / Accepted: 03.12.2022, Online Yayınlanma / Published Online: 25.04.2023)

Anahtar Kelimeler

Maskeli yüz tanıma,
Biyometrik kimlik tanıma,
Otomatik maske giydirmeye,
Görüntü morfleme,
Derin öğrenme

Öz: Covid-19 gibi virüslerin ciddi solunum yolu hastalıklarına neden olması, yüz maskelerinin kullanımını önemli hale getirmiştir. Bu nedenle yüzden kişi doğrulama ve tanıma yapan sistemlerin maskeli yüzler üzerinde de çalışabilmesi beklenmektedir. Yüz tanıma sistemleri için oluşturulan veri setlerinde maskeli yüzler olmamakla birlikte günümüzde farklı model ve desenlerde maskeler kullanılabilir. Bunlar yüz tanıma sistemlerinin başarısını düşürmektedir. Bu çalışmada öncelikle maskeli yüz veri seti üretmek için mevcut yüz veri setlerine farklı tipteki maskelerin giydirilmesine çalışılmıştır. Morfleme tekniği kullanılarak yüzün pozisyonlarına uygun olarak doğala yakın otomatik giydirmeye işlemi gerçekleştirilmiştir. Daha sonra maskeli/maskesiz yüzlerin tanınması için derin öğrenmeye dayalı bir model geliştirilmiş ve otomatik maske giydirmeye tekniği ile oluşturulan veri seti denenmiştir. CASIA-WebFace ve LFW (Labeled Faces in the Wild) veri setleri kullanılarak gerçekleştirilen deneylerde %96.5'in üzerinde maskeli yüz tanıma başarımları elde edilmiştir.

Mask Fitting on Face Images Based on Morphing and Masked Face Recognition

Keywords

Masked face recognition,
Biometric identification,
Automatic mask fitting,
Image morphing,
Deep learning

Abstract: Viruses such as Covid-19 cause serious respiratory diseases, making the use of face masks important. For this reason, it is expected that face recognition and verification systems can also work with masked faces. There are no masked faces in the data sets created for face recognition systems, and today, masks of different models and patterns can be used in daily life. These reduce the success of face recognition systems. In this study, first, a wide masked face data set was produced by fitting different types of masks to the images in the existing face datasets. In the production of masked faces close to nature, the automatic mask fitting process was carried out with the image morphing technique. Then, a deep learning based model was developed for the recognition of masked/unmasked faces, and the dataset created with the automatic mask fitting technique was used to train the model. In the experiments using CASIA-WebFace and LFW (Labeled Faces in the Wild) datasets, masked face recognition performance was achieved over 96.5%.

1. Giriş

Günümüzde biyometrik sistemlerde, yüzde yüze yakın sonuçlar üreten özellikler kullanılabilir hale gelmiştir. Ancak bu özellikler her zaman ulaşılabilir değildir. Doğrulama sistemlerinde kişinin rızası talep edilebilirken, asayiş/gözetleme gibi sistemlerde uzaktan tanıma yapmak gerekmektedir. Covid-19 gibi pandemilerde ise hijyenden ötürü uzaktan ve temassız

algılanabilen biyometrik özellikler daha çok ön plana çıkmaktadır.

Yüz biyometrik özelliği yüksek başarımlı temassız tanıma sistemlerinde kullanılmaktadır. Ancak solunum yollarından bulaşan hastalıklardan ötürü artık yüzde maske kullanımı çoğu zaman gerekli olmakta ve kişilerden maske çıkarmaları sağlık nedeni ile istenemeyecek duruma gelmektedir. Dolayısı ile maskeli yüzlerden kişilerin tanınması veya

*İlgili yazar: gunaya@ktu.edu.tr

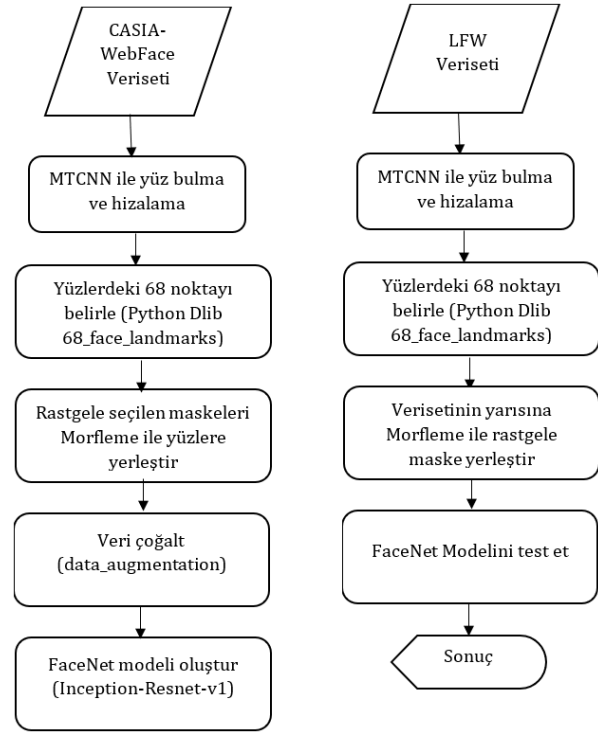
doğrulanması gerekmektedir. Bu kapsamda maskeli yüzler kullanılarak, maske takılıp takılmadığının belirlenmesi ve maske dışında kalan göz ve alın bölgelerinden yüzün tanınması konularında çalışmalar yapılmaktadır.

Maske kullanımının artması ile çeşitli desenlerde maskeler üretilmektedir. Bunlar yüz desenli olabileceği gibi yanlış biyometrik özellik üretimine neden olabilecek türden de olabilmektedir. Bu nedenle maskeli yüzlerin belirlenmesi gerekmektedir. Diğer yandan maske dışında kalan yüz bölgelerinden yüz tanıma işleminde, kullanılacak biyometrik özellikler azaldığından tanıma başarımları düşmektedir.

Bu alanda diğer bir eksiklik ise, yüz tanıma için oluşturulmuş, maskeli yüzleri içeren geniş kapsamlı gerçek bir veri setinin henüz bulunmamasıdır. Bu nedenle bu çalışmada mevcut yüz veri setlerinden maskeli yüz verilerinin oluşturulmasına ve yeni yaklaşımlar kullanılarak daha az biyometrik özellik ile yüz tanıma işleminin gerçekleştirilmesine çalışılmıştır.

Bu çalışmada, ilk olarak morfleme tekniği kullanılarak farklı tip maskelerin farklı pozlardaki yüzlere oturtulmasına çalışılmıştır. Bunun için CASIA-WebFace veri setindeki yüzler kullanılmıştır. Morfleme tekniği ile daha çok nokta kullanıldığı için maskeler doğala yakın bir şekilde yüzlere oturtulmaktadır. Maske oturtma işlemi için yüz üzerindeki 68 temel özellik noktası, DLIB [1] tabanlı bir yüz detektörü [2] ile tespit edilmiştir. Bu noktalardan yüzün alt kısmını (burun ağız ve çeneyi çevreleyen) temsil eden noktalar kullanılarak, görüntü morfleme tekniğine dayalı olarak, yüze maske oturtulmuştur. Çalışmada daha sonra yüz tanıma için bir FaceNet modeli oluşturulmuştur [3]. Bu modelde, Inception-ResNet-v1 [4] derin ağ mimarisi kullanılmıştır. Model eğitimi kolaylaştırmak amacıyla MTCNN [5] (Multi Task Cascaded Convolutional Neural Network) modeli ile yüz bölgesi belirlenerek gereksiz verilerin atılması ve ağın yalnızca anlamlı bilgileri işleme sağlanmıştır. Son olarak morfleme tekniğine dayalı olarak üretilen maskeli yüz görüntüleri ve maskesiz yüz görüntüleri birlikte kullanılarak oluşturulan veri seti üzerinde modelin başarımları değerlendirilmiştir. Sistemin genel akış şeması Şekil 1'de verilmiştir.

Maskeli yüz tanıma alanında yapılan çalışmalar iki gruba ayrılabilir. Birinci grup, maskeli yüzlerden, maskenin kapatmadığı yüz bölgesini (göz ve alın) belirleyip bu bölge üzerinden kişiyi tanımayla çalışmaktadır. [6, 7, 9-11]. Qi ve Yang, maskeli



Şekil 1. Sistemin genel akış şeması

görüntülerden MTCNN ile yüz bölgesini belirlemiş, maskesiz kısımdan yerel ikili örüntülerle öznelik çıkarmış ve destek vektör makineleri (SVM) ile sınıflandırmıştır [6]. Deneyler WIDER FACE veri setinden alınan 3894 görüntü ve MAFA veri setinden alınan 4064 görüntü üzerinde gerçekleştirilmiştir. Alrikabi ve Alibreheemi çalışmalarında, GoogleNet ve VGGNet-16'dan çıkardıkları öznelikleri birleştirmiş ve SVM ile sınıflandırmıştır [7]. Görüntülerdeki yüzler MTCNN ile tespit edilmiş, maskesiz kısım kesilmiş ve ağlara giriş olarak verilmiştir. Deneylerde RMFRD ve SMFRD veri setleri kullanılmıştır [8]. Mazli Shahar ve Mazalan yüz görüntülerindeki 68 kontrol noktasını belirledikten sonra bu noktalardan göz burun çevresini ifade eden 31 tanesi kullanarak maskeli yüz tanıma yöntemi önermiştir [9]. Hariri maskeli yüzlerin göz ve alın bölgesini kullanarak yüz tanıma gerçekleştirmiştir [10]. VGGNet-16 ağı ile çıkarılan özneliklerin sınıflandırılmasında çok katmanlı algılayıcı (perceptron) kullanılmıştır. Deneylerde ise RMFRD veri seti kullanılmıştır. Vu vd. de maskeli yüzlerden maskesiz kısmı çıkarıp yüz tanıma işlemi gerçekleştirmiştir [11].

Maskeli yüz tanıma alanındaki diğer çalışmalar ise, kişilerin maskesiz ve maskeli görüntüleriyle sistemi eğitip, iki görüntünün aynı ya da farklı kişiye ait olduğuna karar vermeye çalışmaktadır [12, 13]. Ejaz ve Islam MTCNN ile belirledikleri yüz görüntülerinden Google FaceNet modeli ile yüz özelliklerini çıkarmış ve SVM ile sınıflandırmıştır [12]. Jaween vd. ise yüz tanıma başarılı derin öğrenme sistemlerinin maskeli yüz tanımadaki başarımlarını incelemiştir. FaceNet, VGGNet, LeNet, MobileFaceNet gibi modellerini maskeli-maskesiz, maskeli-maskeli ve maskesiz-maskesiz yüz çiftleriyle denemiştir [13].

Gerçek maskeli yüz görüntülerinden oluşan veri setinin yanında, sentetik olarak üretilmiş maskeli yüzlerden oluşan bir veri seti de deneylerde kullanılmıştır.

Bunların dışında, maskeli yüzlerden maskesiz yüzleri üretip tanıma yapan çalışmalar da bulunmaktadır [14, 15]. Luo vd. çekişmeli üretici ağları (generative adversarial network-GAN) kullanarak, EyesGAN adını verdikleri bir model ile maskesiz yüzleri sentezlemeye çalışmıştır [14]. Malakar vd. ise maskeli yüzlerden maskesiz yüzlerin üretiminde temel bileşen analizini kullanmayı önermiştir [15].

Mevcut veri setlerinde maskeli yüz görüntü sayısı azdır. Bu nedenle maskeli yüzlerin sentetik olarak üretilmesine çalışılmaktadır. Anwar ve Racchowdhury yüz görüntülerine maske yerleştirmek için 6 anahtar nokta konumunu kullanmıştır [16]. Yüzün dönme açısına göre 3 aşadaki çeşitli tipte maske görüntüleri yüz görüntülerine eklenmiştir. Wang vd. ise çalışmalarında 3 çeşit veri seti oluşturmuştur [8]. Bunlar maskeli yüz tespiti veri seti (MFDD), sentetik maskeli Yüz veri seti (SMFRD) ve gerçek maskeli yüz veri seti (RMFRD) olarak adlandırılmıştır. Çalışmada maskenin nasıl yerleştirildiği konusunda bilgi verilmemiştir. Mevcut çalışmalarda, sentetik olarak maskeli yüz görüntülerinin üretilmesinde, maskenin yüze tam oturmaması sorunu ortaya çıkmaktadır.

Yazının organizasyonu: Bölüm 2’te önerilen yöntem detaylı bir şekilde açıklanmıştır. Deneyler ve bulgular 3. Bölümde verilmiş ve sonuçlar Bölüm 4’te tartışılmıştır.

2. Materyal ve Metot

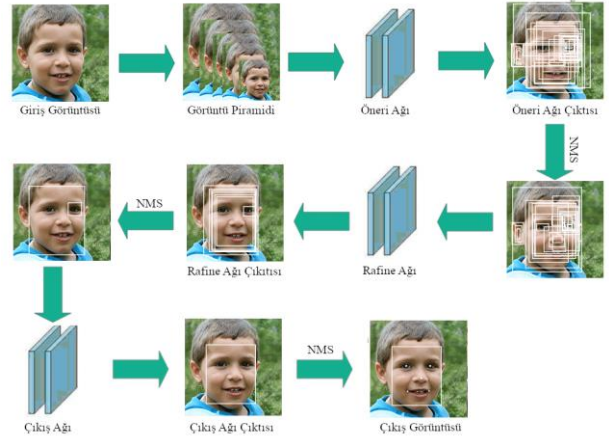
Bu çalışma iki aşamadan oluşmaktadır. İlk aşamada sentetik maskeli yüz görüntüsü üretmek için, maskenin yüze daha iyi oturmasını sağlayan, morflemeye dayalı bir yöntem önerilmiştir. Sonraki aşamada, derin öğrenme modeli ile yüz tanıma sistemi oluşturulmuş ve üretilen maskeli yüz görüntüleri ile denenmiştir.

2.1. Maskeli Yüz Görüntüsü Üretimi

Maskeli yüzlerin üretilmesinde öncelikle görüntülerde yüz algılanmış ve hizalanmıştır. Daha sonra yüzdeki kontrol noktaları belirlenerek ilgili bölgeye görüntü morfleme yöntemi ile maske yerleştirilmiştir.

2.1.1. Yüz Algılama ve Hizalama

Yüz algılama ve hizalama işlemini gerçekleştirmek için FaceSSD [17], Dlib [1], OpenCV [18], MTCNN [5] gibi çeşitli algoritmalar geliştirilmiştir. Bu çalışmada yüz algılama ve hizalama için MTCNN modeli kullanılmıştır. MTCNN, 3 aşamalı sinir ağlarından oluşan modern bir yüz algılama aracıdır. Bu aşamalar Şekil 2’de görülmektedir.



Şekil 2. MTCNN aşamaları

MTCNN yüz belirleme mimarisinde ilk olarak, farklı boyutlardaki yüzleri algılamak için görüntü birden çok kez yeniden boyutlandırılır. Ardından P-ağı (Proposal/Öneri), görüntüleri tarayarak ilk algılamayı gerçekleştirir. Burada düşük bir eşik değeri kullanıldığı için birçok yanlış pozitifte algılanır. Önerilen bölgeler (birçok yanlış pozitif içeren), hassas sınırlayıcı kutular elde etmek için filtreleme yapan R-ağına (Refine/Arıtma) girdi olarak verilir. Son aşama olan O-ağı (Output/Çıkış) sınırlayıcı kutuların son iyileştirmesini gerçekleştirir. Bu şekilde yalnızca yüzler algılanmakla kalmaz, aynı zamanda sınırlayıcı kutular da doğru ve kesin şekilde oluşturulur ve yüzler belirlenmiş olur. MTCNN güçlü bir algoritma olup, farklı boyut, aydınlatma koşulu ve dönüş açısına sahip görüntülerde yüzleri doğru şekilde algılayıp hizalamaktadır. Şekil 3’te MTCNN modeli ile belirlenen yüz görüntüleri verilmiştir.

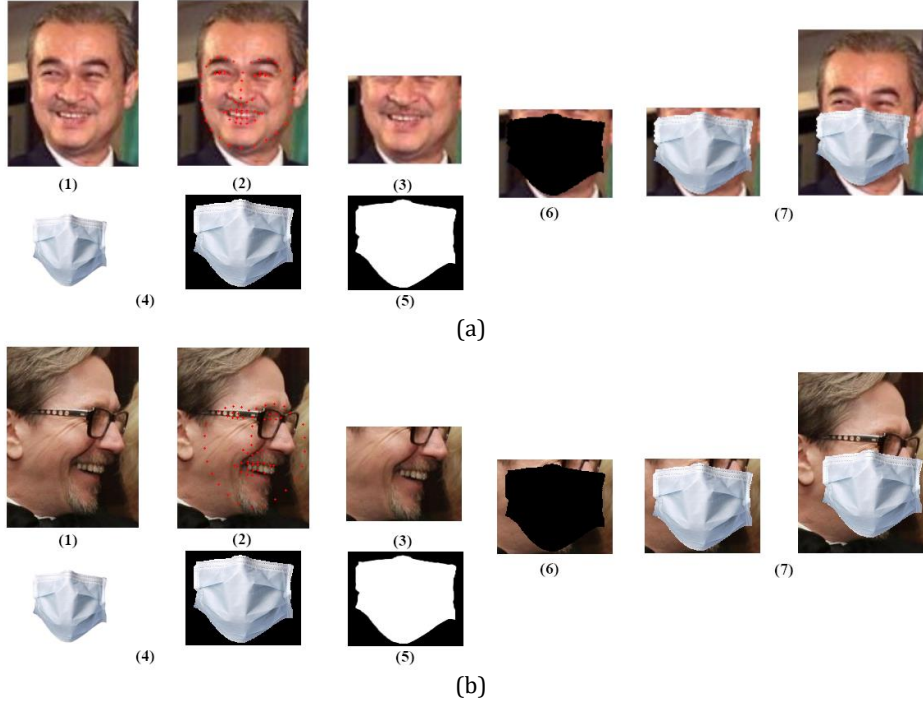


Şekil 3. MTCNN ile yüz algılama ve hizalama örnekleri

2.1.2. Yüz Görüntüsüne Maskenin Yerleştirilmesi

Yüze maske oturtma işlemlerinde maske ile yüzün aynı hizada olmaması, ölçeklemenin iyi yapılamaması, maskenin doğalından çok farklı oturtulmasına sebep olmaktadır. Mevcut çalışmalardaki yüz maskeleme algoritmalarının temel adımları şu şekildedir:

1. Görüntüdeki yüz bölgesi tespit edilir.
2. Yüzdeki 68 kontrol noktası tespit edilir.



Şekil 4. Yüz maskeleme algoritmasının temel adımları

3. Burun ve ağız içeren maske bölgesi belirlenir ve kesilir. Bu alan 49 ile 68 noktalar arasındadır.
4. Maske görüntüsü, belirlenen maske bölgesinin boyutuna göre yeniden boyutlandırılır.
5. Boyutlandırılan maske görüntüsünden ikili maske çıkartılır.
6. İkili maske kullanılarak yüzdeki maske alanı belirlenir.
7. Belirlenen bölgeye maske eklenir ve görüntü güncellenir.

Yüz maskeleme algoritma adımlarının örnek iki görsel üzerine uygulandığı Şekil 4'te gösterilmektedir. Şekil 4-a'da algoritma iyi sonuç verirken Şekil 4-b'de başarısız olmaktadır. Maskenin düzgün oturtulmadığı örnek görüntüler Şekil 5'te verilmiştir.



Şekil 5. Yanlış yerleştirilen maske görüntüleri

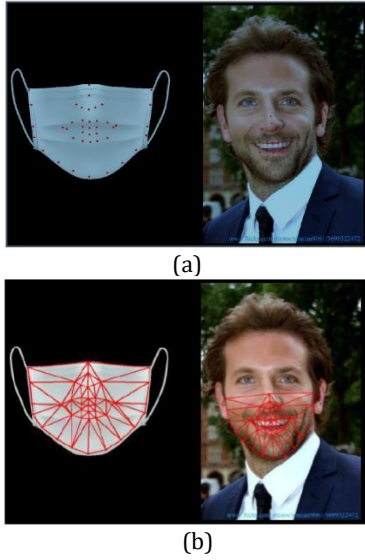
Bu çalışmada ise, maskelerin daha doğru oturmasını sağlayan görüntü morfleme tekniğine dayalı yeni bir yüz maskeleme yaklaşımı önerilmiştir. Önerilen görüntü morflemeye dayalı yüz maskeleme yaklaşımı aşağıdaki aşamalardan oluşmaktadır.

Kontrol Noktalarının Belirlenmesi: Maskeyi yüze oturtmak için ilk aşama ilgili bölgenin kontrol noktalarının tespitidir. Bu amaçla Dlib [1] kütüphanesi ile yüzdeki 68 kontrol noktası tespit edilmiş daha sonra çene, burun ve ağızı çevreleyen kontrol noktaları maske yerleştirme işlemi için kullanılmıştır. Yerleştirilecek maskenin bu bölgelere karşılık gelen aynı sayıdaki kontrol noktaları ise manuel olarak işaretlenmiştir (Şekil 6-a).

Üçgensel Bölgelerin Oluşturulması: Bu aşamada, önceki aşamada elde edilen maske kontrol noktaları kullanılarak birbirleriyle bağlantılı üçgensel bölgeler elde edilir. Yüz görüntüsünde de her bir üçgeni oluşturan noktalar seçilerek, maskedekilere karşılık gelen üçgensel bölgeler üretilir. Çalışmada Delaunay üçgenleme yönteminden faydalanılmıştır (Şekil 6-b)

Üçgensel Bölgelerin Eğriltilmesi: Maske ve yüz görüntülerindeki kontrol noktalarının pozisyonları birbirinden farklıdır. Maskeyi düzgün bir şekilde yüze oturtabilmek için, maske görüntüsündeki üçgensel bölgelerin, yüz görüntüsünde karşılık gelen üçgensel bölgenin şekline eğriltilmesi (warping) gerekmektedir. Birbirine karşılık gelen üçgensel bölgelerin tanımlanmasından sonra, morfleme işlemi geometrik dönüşüm problemi haline gelmektedir. Bu noktada afin dönüşüm kullanılmaktadır. Yüz görüntüsüne göre eğriltmiş maske görüntüsü Şekil 7-a'da görülmektedir.

Hedef Görüntünün Oluşturulması: Bu aşamada yüz görüntüsünde ilgilenen bölge silinir (Şekil 7-b). Son olarak, eğriltme işlemi ile yeniden oluşturulan maske görüntüsü yüz görüntüsüne eklenir (Şekil 7-c).



Şekil 6. a) Maske ve yüz kontrol noktaları. b) Maske ve yüz üçgensel bölgeleri.



Şekil 7. a) Yüz görüntüsüne göre eğriltilmiş maske b) Maske eklenecek yüz bölgesi c) Maskeli yüz görüntüsü.

2.2. Derin Öğrenme

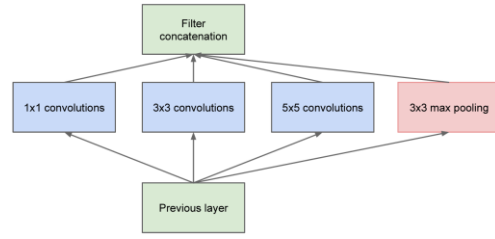
Makine öğrenmesi yaklaşımlarından olan derin öğrenme, bulut veya GPU'lar kullanılarak günümüzde bilgisayar bilimcilerin kümeleme veya sınıflandırma çalışmalarında sıklıkla tercih ettiği, yapay sinir ağlarına göre çok daha fazla katman ve aşamadan oluşan yapılardır. Bunlardan Evrişimsel sinir ağı (ESA) mimarileri, imge sınıflama, nesne algılama, tanıma ve tespit işlemlerinde başarılı sonuçlar üretmektedir. ESA, yapay sinir ağlarından esinlenilerek geliştirilen, verileri uçtan uca öğrenen bir derin öğrenme mimarisidir. Büyük ölçekli verilerle çalışabilen ESA, uzun eğitim süreci ve yerel minimumlara takılabilme problemlerine sahiptir. Bu çalışmada kullanılan yöntemlerin belirli özellikleri aşağıda verilmiştir.

2.2.1. GoogleNet Ağı

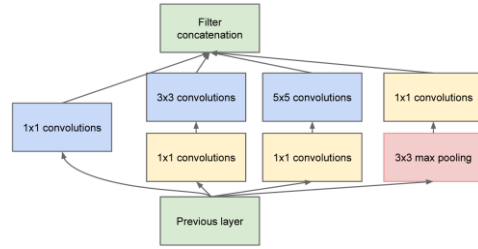
GoogleNet ağı, hesaplama maliyetini artırmadan daha geniş ve daha derin bir model olarak 2014 yılında geliştirilmiştir. 22 adet katmandan oluşan ağa Inception adı verilmiştir. Temel Inception yapısı, Şekil 8'de görüldüğü gibi, aynı girdiye havuzlama ve birden fazla evrişim işleminin uygulanması esasına dayanmaktadır. Inception modelinde, girdi maksimum havuzlama ve 1×1 , 3×3 ve 5×5 boyutlarında filtrelemeden geçirilerek evrişim işlemine tabi tutulur. Bu işlemlerin sonuçları birleştirilerek tek bir sonuç

elde edilir. Bu sayede GoogleNet ağı, genel ve spesifik özellikleri aynı anda çıkarabilmektedir [19].

Inception modelinin 5×5 evrişiminde hesaplama maliyetinin artması en büyük problemlerdendir. Ayrıca havuzlama katmanlarının dahil olması ile hesaplama maliyeti oldukça artmaktadır. Her bir aşamada bu evrişim katmanının çıktısı ile havuzlama katmanının çıktısının birleştirilmesi çıktı boyutunu çok fazla büyütülmektedir [19]. Bu nedenle her işlemden önce boyut düşürülmelidir. Boyut düşürme işlemi Şekil 9'da verildiği gibi, 5×5 ve 3×3 filtrelerinin uygulanmasından önce 1×1 boyutlarındaki filtre uygulaması ile gerçekleştirilmektedir [19].

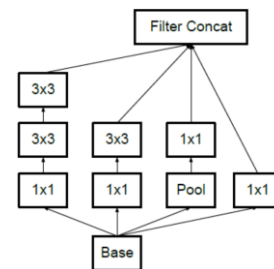


Şekil 8. Basit Inception yapısı



Şekil 9. Boyut düşürme işleminden sonra Inception yapısı

Inception modeline daha sonradan paket normalizasyonu eklenerek Inception V2 modeli oluşturulmuştur. Paket normalizasyonu, tüm katmanların ardında, çıktı verilerinin ortalama ve standart sapmasına dayalı olarak katman çıkışlarındaki verilerin normalizasyonu şeklinde gerçekleştirilmektedir. Bu işlem sayesinde, tüm sinir ağı, aynı aralıkta sıfır ortalamalı sonuçlar üretmiş olur [20]. GoogleNet ağının sonraki diğer modeli Inception V3'te ise, ağın genişlik ve derinliğinin optimizasyonu ile ağdan maksimum bilgi akışının sağlanmasına çalışılmıştır. Ağda, genişlik, derinlik ile doğru orantılı olarak artmaktadır. Inception V3 ağında, 5×5 ve 7×7 boyutlarında filtreler, iki ya da üç tane 3×3 boyutlu filtrelerle değiştirilerek (Şekil 10) optimize edilmeye çalışılmaktadır [21].

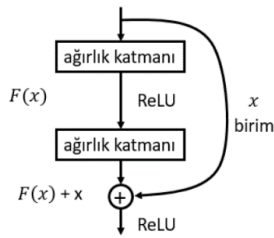


Şekil 10. Inception V3 ağı

Son olarak Inception V4 modelinde, Resnet ağının gelişimiyle, daha hızlı eğitim yapılmasını sağlamak amacıyla, Inception modeli ile kalıntı (residual) ağ bağlantıları birleştirilerek geliştirilmiştir [4].

2.2.2. Microsoft Resnet Ağı

Geleneksel ağlardan farklı yapıya sahip Resnet ağı, bazı kısa yolların klasik ağa eklenmesiyle elde edilen kalıntı bloklardan oluşmaktadır. Bu kalıntı blokta, x girdisi evrişim->aktivasyon->evrişim serisinden geçirilerek, $f(x)$ fonksiyonuna dönüştürülmektedir. Sonraki adımda, $f(x)$ fonksiyonuna x giriş değeri eklenerek $h(x) = f(x) + x$ fonksiyonu üretilmektedir. Klasik evrişim işleminde $h(x)=f(x)$ iken bu ağda, girdiye evrişim işlemi uygulandıktan sonra orijinal veri eklenmiştir [22]. Ağın blok diyagramı Şekil 11'de gösterilmiştir.



Şekil 11. Kalıntı (Residual) öğrenme bloğu

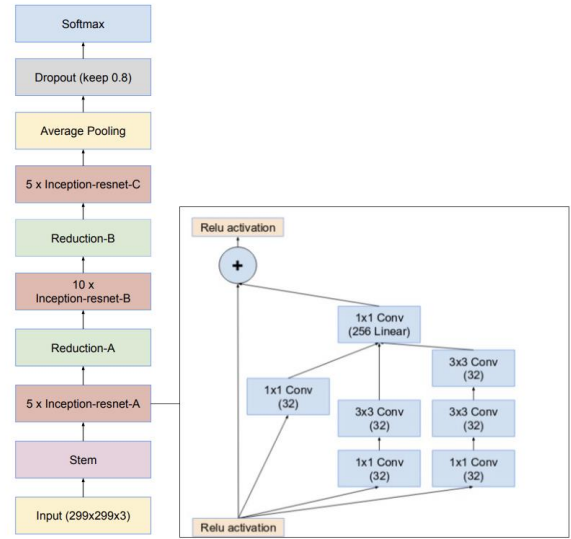
2.2.3. Inception-ResNet-v1 Ağı

Inception yapısıyla kalıntı (residual) ağ bağlantıları birleştirilerek Inception-ResNet-v1 elde edilmektedir [22]. Bu ağ önceki ağların özelliklerini taşımaktadır. Ağın genel yapısı Şekil 12'de gösterilmiştir.

Çalışmada FaceNet modeli oluşturulurken Inception-ResNet-v1 derin ağ mimarisi kullanılmıştır. Bu ağ ile yüz görüntülerinden öznelikler çıkarılmış ve yüz tanıma için kullanılmıştır. Sonraki alt bölümde tasarlanan FaceNet modeli açıklanmıştır.

2.3. FaceNet modeli

FaceNet [3] modeli, yüz tanıma, doğrulama ve kümeleme işlemlerinde oldukça iyi sonuçlar elde



Şekil 12. Inception-ResNet-v1 yapısı

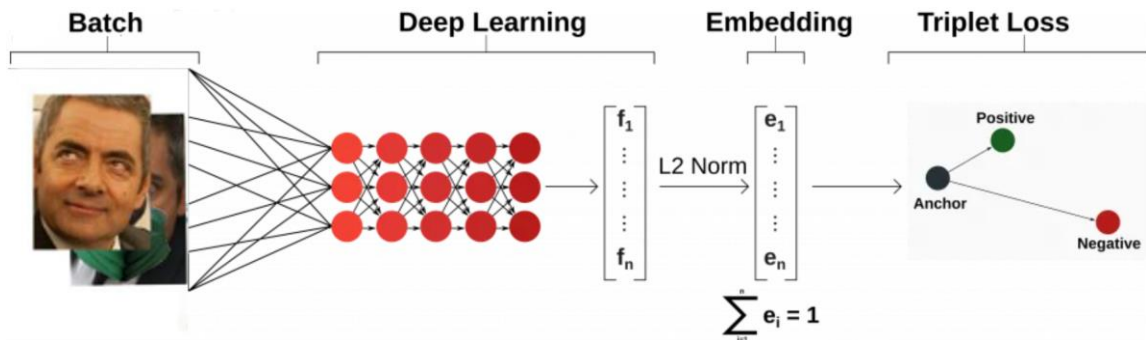
edilmiş bir mimaridir. FaceNet modelinde, çok boyutlu girdi olarak alınan yüz resimleri 128 boyutuna sahip bir özellik vektörüne dönüştürülür. Bu özellik vektörlerinin birbirine olan benzerliği hesaplanır. Benzerlik oranı değeri yüksek olan özellik vektörleri birbirine daha çok benzer olarak sınıflandırılır. FaceNet modelinde, özellik vektörünün üretilebilmesi için üçlü kayıp (triplet loss) fonksiyonu ve evrişimli sinir ağları (CNN) birlikte kullanılır. Şekil 13'te görülen FaceNet modelinin mimarisi aşağıdaki bileşenlerden oluşmaktadır.

Giriş katmanı: (112, 112, 3) boyutlu görüntüler

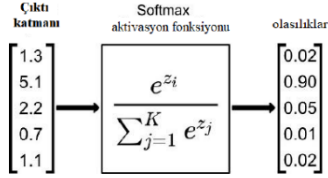
Ağın derin mimarisi (gizli katmanları): Bunu tasarlamak için Inception-Resnet-v1 [22] kullanılmıştır. Tahmin doğruluğunu artırmak için, eğitim aşamasında bu katmanların ağırlık değerleri ayarlanacaktır.

Özellik vektörü (Embedding): Derin mimarisinin son katmanından gelen değerler L2 normalizasyonuna tabi tutulur ve bu katmanın girdisi olarak kabul edilir. Bu katman daha sonra test için kullanılacaktır.

Çıkış katmanı: Bu katman derin mimariye doğrudan bağlıdır. Softmax aktivasyon fonksiyonu ile normalize edilir ve tahmini sınıf elde edilir (Şekil 14).



Şekil 13. FaceNet modelinin mimarisi



Şekil 14. Softmax aktivasyon fonksiyonu

FaceNet mimarisinde kullanılan üçlü kayıp fonksiyonunun hesaplanmasında, veri kümesinden birbirinden farklı üçlü gruplar seçilir. Üçlü kayıp değeri seçilen 3 imge üzerinden hesaplanır. Seçilen üç imgenin biri referans (anchor) imge olacaktır. Diğer imgelerden biri referans imgeye benzeyen (pozitif) ve diğeri benzemeyen (negatif) imgelerdir. FaceNet modelinde kayıp değeri; referans imge ile pozitif imge arasındaki benzerlik arttıkça küçülür. Benzer şekilde, referans imge ile negatif imge arasındaki benzerlik artışı kayıp değerinin artmasına neden olur. Kayıp değeri kullanılarak ağırlıkların güncellenmesi ile tamamlanan eğitim sonunda FaceNet, girdi olarak aldığı imgeyi en iyi temsil eden özellik vektörünü çıkarmayı öğrenmiş olur.

Saha uygulamalarında, iki yüzün aynı kişiye ait olup olmadığını tespit etmek için, girdi olarak alınan yüz imgesi ile veri kümesindeki yüz imgelerinin FaceNet modeli ile özellik vektörleri üretilir. Daha sonra, uzaklık metrikleri kullanılarak bu özellik vektörlerinin benzerliği ölçülür. Benzerlik değeri önceden belirlenen deneyimlerle kabul görmüş eşik değerinin altında ise doğrulamanın başarılı olduğu, eşik değerinin üstünde ise yüz imgelerinin farklı kişilere ait olduğu şeklinde yorumlanır.

3. Bulgular

3.1. Veri Setleri

3.1.1. CASIA-WebFace veri seti

CASIA-WebFace veri seti [23] yüz doğrulama ve yüz tanıma işlemlerinde kullanılmaktadır. Veri seti internetten toplanan 10.575 gerçek kişiye ait 494.414 yüz görüntüsünden oluşmaktadır. Bu görüntüler sınıflar arasında homojen olarak dağılmamış olup çoğu sınıf 0 ile 100 arasında görüntü içermekte, bazı sınıflarda ise 800 görüntü bulunmaktadır. Bu problemi çözmek ve homojen veri seti oluşturmak için her sınıftan sabit miktarda görüntü seçilmiştir. Çalışmada tüm verilere yüz hizalama uygulandıktan sonra her sınıftan rastgele 10 görüntü seçilmiş, 10575 sınıf için yaklaşık 100.000 görüntü elde edilmiştir. Daha sonra oluşturulan veri setindeki görüntülerden maskeli yüzler üretilmiş ve veri çoğaltma işlemi (gürültü ekleme, döndürme, parlaklık değiştirme) uygulanarak toplamda 400.000 adet görüntü elde edilmiştir. Bu veri seti oluşturulan FaceNet modelinin eğitiminde kullanılmıştır. CASIA-WebFace veri setinden örnek görüntüler, önerilen yöntemle üretilen



(a)



(b)

Şekil 15. CASIA-WebFace veri seti a) Orijinal yüz görüntüleri ve üretilen maskeli yüzler b) Veri çoğaltma örnekleri

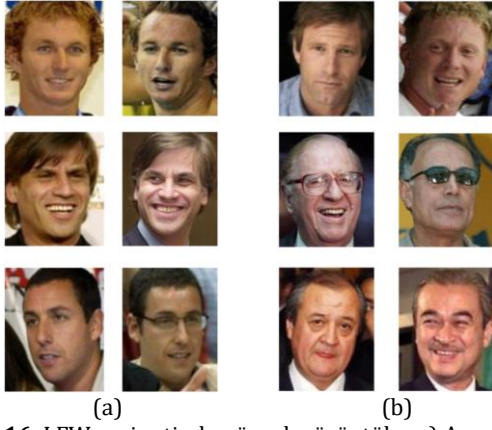
maskeli yüzler ve veri çoğaltma örnekleri Şekil 15'te verilmiştir.

3.1.2. LFW veri seti

LFW (Labeled Faces in the Wild) veri seti [24] internetten toplanmış 13.233 yüz görüntüsü içermektedir. Veri seti 1.680 kişinin bir ya da daha fazla görüntüsünden meydana gelmektedir. Standart LFW değerlendirme protokolünde doğrulama başarımları 6.000 yüz görüntüsü çifti üzerinden hesaplanmaktadır. Çalışmada LFW veri seti, model performansını ve doğruluğunu test etmek için kullanılmıştır. Bu amaçla ikişer ikişer test örnekleri hazırlanarak 6000 test durumu oluşturulmuştur. Bunlardan 3000 test durumu aynı kişiye ait görüntüleri test etmek için, 3000 test durumu farklı kişilerin görüntülerini test etmek için kullanılmıştır. Veri setinden aynı ve farklı kişilere ait örnek çifti görüntüleri Şekil 16'da, önerilen yöntemle üretilen maskeli görüntüler ise Şekil 17'de görülmektedir. Modelin eğitiminde ve testinde kullanılan veri setleri ve örnek sayıları Tablo 1'de verilmiştir.

Tablo 1. Veri setlerindeki görüntü sayıları

	Veri seti	Örnek sayısı
Eğitim	CASIA-WebFace	400000
Test	LFW	12000



Şekil 16. LFW veri setinden örnek görüntüler a) Aynı kişiye ait görüntü çiftleri b) Farklı kişilere ait görüntü çiftleri



Şekil 17. LFW veri setinden orijinal yüz görüntüleri ve üretilen maskeli yüzler

3.2. Uygulama

Çalışmada maskeli yüz tanıma işlemi PyCharm IDE'si ile gerçekleştirilmiştir. Uygulamada OpenCV, Dlib, numpy, Tensorflow Python modülleri kullanılmıştır. Donanımsal anlamda oldukça güçlü gereksinimleri olan uygulama için eğitim verileri Google Colaboratory'de eğitilmiştir. Sunucunun özellikleri, CUDA framework'u destekleyen Nvidia Tesla V100 16GB ekran kartı, tek çekirdekli Intel Xeon'un iki işlemcisi ve 13GB RAM bellek şeklindedir.

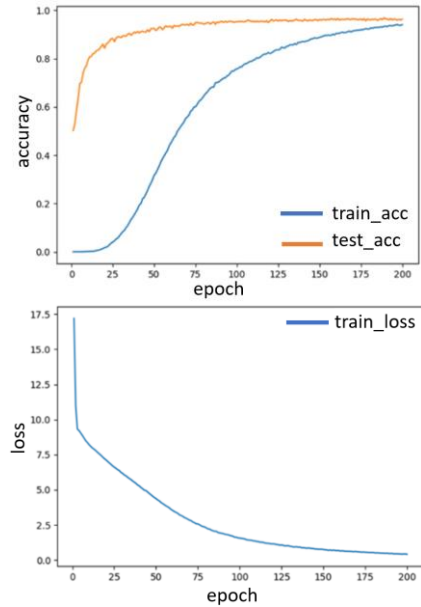
Veri seti hazırlarken yüz hizalamak için MTCNN modeli kullanılmıştır. Maskeli yüz görüntüleri, Python'da Dlib aracını kullanan 68_face_landmarks modülü yardımıyla oluşturulmuştur. Bu modül giriş görsellerindeki yüzleri algulamakta ve özellik noktalarını çıkartmaktadır. Özellik noktaları çıkartılan yüzlere maskeler görüntü morfleme tekniğiyle eklenmiştir. Veri setindeki görüntü sayısını artırmak için görüntü işleme teknikleri yardımıyla veri çoğaltma (data augmentation) işlemi uygulanmıştır. Inception-Resnet-v1 mimarisi kullanılarak bir FaceNet modeli oluşturulmuştur. Ağırlık katsayılarını iyileştirmek için optimize edici olarak "adam optimizer" kullanılmıştır. Özellik vektörü boyutu en iyi doğrulama değerine sahip olan 128 seçilmiştir [3]. Kayıp değeri ve doğruluk oranı hesaplamak için "cross entropy" kullanılmıştır.

Oluşturulan model, üretilen CASIA-WebFace veri seti ile eğitilmiş, LFW veri seti ile değerlendirilmiştir. LFW veri setindeki 6000 test durumu için (3000 test aynı kişi, 3000 test farklı kişi), iki görüntü aynı kişiye aitse

birbirine çok yakın olmalı ve hesaplanan mesafe küçük olmalıdır. Farklı kişilerin görüntüleri ise birbirinden çok uzak olmalıdır. Burada bir eşik listesi tanımlanmış, bu liste 0 ile 4 arasında dağıtılmıştır. Görüntüler arasındaki mesafe hesaplandıktan sonra eşik değerine göre karar verilmektedir. En iyi test doğruluğu sağlayan eşik değeri en iyi eşik değeri olarak seçilmiştir. Mesafe hesaplamak için Öklid uzaklık ölçütü kullanılmıştır.

3.3. Deneysel Bulgular

Oluşturulan FaceNet modelinin 10575 sınıfa ait yaklaşık 400000 maskeli ve maskesiz görüntüyle 200 eğitim turu (epoch) boyunca eğitilmesi ve 6000 test durumuyla doğrulanması sonucunda elde edilen başarımlar Şekil 18'de gösterilmektedir. Grafiklerde eğitim tur sayısının ilerlemesi ile modelin hata değerlerinin düştüğü ve doğruluğunun arttığı görülmektedir.



Şekil 18. Eğitim ve test sonuçları

Tanıma sistemlerinin performansının ölçülmesinde kullanılan tanıma doğruluğu (accuracy) karışıklık matrisindeki (confusion matrix) değerler üzerinden aşağıdaki denklemle hesaplanır.

$$\text{Doğruluk} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Burada TP (true positives) doğru pozitifleri, TN (true negatives) doğru negatifleri, FP (false positive) yanlış pozitifleri, FN (false negative) ise yanlış negatifleri ifade etmektedir.

Modelin eğitimi ile deneyimlenen test kümesindeki en iyi eşik değeri 1.31 olarak belirlenmiştir. Bu eşik değeri kullanılarak üretilen Karışıklık Matrisi (Confusion Matrix) Şekil 19'da görülmektedir. Karmaşıklık matrisine göre aşağıdaki bilgiler elde edilmiştir.

1. İki görüntünün aynı kişiye ait olduğu tahmin edilmiştir, gerçekte bu doğrudur (TP): 2884
2. İki görüntünün aynı kişiye ait olduğu tahmin edilmiştir, gerçekte bu yanlıştır (FP): 93
3. İki görüntünün farklı kişilere ait olduğu tahmin edilmiştir, gerçekte bu doğrudur (TN): 2907
4. İki görüntünün farklı kişilere ait olduğunu tahmin edilmiştir, gerçekte bu yanlıştır (FN): 116

Gerçek Değerler

		Gerçek Değerler	
		Pozitif	Negatif
Tahmin edilen Değerler	Pozitif	2884 (TP)	93 (FP)
	Negatif	116 (FN)	2907 (TN)

Şekil 19. Modelin karmaşıklık matrisi

Bu sonuçlara göre, oluşturulan modelin test doğruluğu Tablo 2'de görüldüğü gibi %96.5 olarak hesaplanmıştır.

Tablo 2. Önerilen modelin başarımı

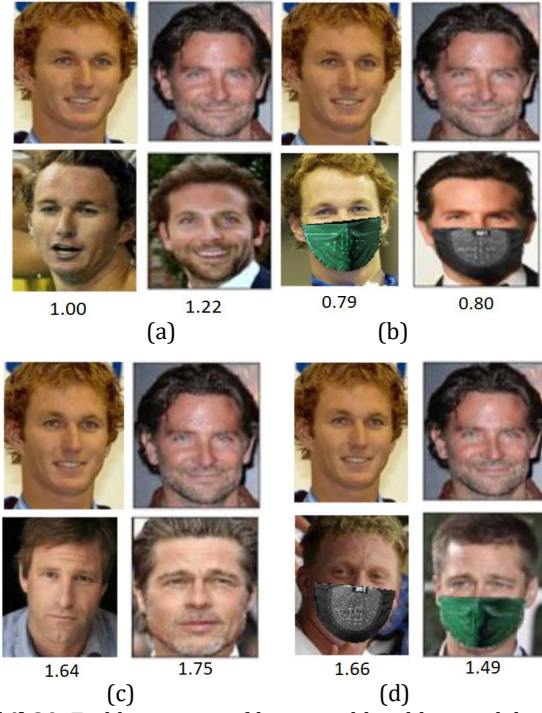
Eğitim Seti	Test Seti	En İyi Eşik Değeri	Başarım
CASIA-WebFace	LFW	1.3	%96.5

Modelin farklı görüntü çiftleri üzerinde çalıştırılması ile hesaplanan özellik vektörleri arasındaki Öklid uzaklıkları hesaplanmış ve elde edilen sonuçlar Şekil 20'de verilmiştir. Şekilde görüldüğü gibi iki kişinin görüntüsü temel olarak alınmış, aynı kişilerin maskesiz/maskeli görüntüleri ve farklı kişilerin maskesiz/maskeli görüntüleri kullanılarak model sonucu hesaplanmıştır.

Sonuçlara göre aynı kişinin görüntüleri (Şekil 20-a, Şekil 20-b) arasındaki mesafe farklı kişi görüntüleri (Şekil 20-c, Şekil 20-d) arasındaki mesafeden daha küçüktür ve bu mesafe test kümesinden elde edilen en iyi eşik değeri olarak hesaplanan 1.31'i aşmamıştır.

4. Tartışma ve Sonuç

Bu çalışmada mevcut yüz tanıma sistemlerinin maske kullanımını nedeniyle başarısız hale gelmesini engellemek amacıyla, bu sistemlerde kullanılan mevcut veri setlerindeki görüntülerden maskeli yüzlerin elde edilmesi için yeni bir yaklaşım önerilmiştir. Daha sonra bir FaceNet modeli tasarlanmış, modelin eğitiminde CASIA-WebFace veri setindeki maskesiz ve çalışmada üretilen maskeli yüzler kullanılmıştır. Model LFW veri seti ile değerlendirilmiştir. Elde edilen sonuçlar yapay bir şekilde üretilen maske görüntüleri ile başarılı bir yüz tanıma gerçekleştiğini göstermektedir.



Şekil 20. Farklı görüntü çiftleri için elde edilen model sonuçları a) aynı kişinin görüntüleri b) aynı kişinin maskeli görüntüleri c) farklı kişi görüntüleri d) farklı maskeli kişi görüntüleri

Etik Beyanı

Bu çalışmada, "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması gerekli tüm kurallara uyulduğunu, bahsi geçen yönergenin "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbirinin gerçekleştirilmediğini taahhüt ederiz.

Kaynakça

- [1] Dlib C++ Kütüphanesi, 2021. <http://dlib.net/> (Erişim Tarihi : 15.12.2021).
- [2] Yüz Hizalama Detektörü GTX kütüphanesi, 2015. <https://github.com/davisking/dlib-models> (Erişim Tarihi : 22.09.2021).
- [3] Schroff, F., Kalenichenko, D., Philbin, J. 2015. Facenet: A Unified Embedding for Face Recognition and Clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7-12 June, Boston, MA, USA, 815-823.
- [4] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A. A. 2017. Inception-v4, Inception-Resnet and The Impact of Residual Connections on Learning. Thirty-first AAAI Conference on Artificial Intelligence (AAAI17), 4-9 February, San Francisco, California USA, 4278-4284. 2017.
- [5] Zhang, K., Zhang, Z., Li, Z., Qiao, Y. 2016. Joint face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23(10), 1499 – 1503.

- [6] Qi, C., Yang, L. 2020. Face Recognition in The Scene of Wearing A Mask. 2020 International Conference on Advance in Ambient Computing and Intelligence (ICAACI). 12-13 September, Ottawa, ON, Canada, 77-80.
- [7] Alrikabi, J.M., Alibraheemi, K. H. 2021. A Combination Approach for Masked Face Recognition Based on Deep Learning. *AM*, 9(3), 499-520.
- [8] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., Yi, P., Jiang, K., Wang, N., Pei, Y., Chen, H., Miao, Y., Huang, Z., Liang, J. 2020. Masked Face Recognition Dataset and Application. arXiv preprint arXiv:2003.09093.
- [9] Mazli Shahar M.S., Mazalan, L. 2021. Face Identity for Face Mask Recognition System. 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). 3-4 April, Penang, Malaysia, 42-47.
- [10] Hariri, W. 2021. Efficient Masked Face Recognition Method during the COVID-19 Pandemic. arXiv preprint arXiv:2105.03026.
- [11] Vu, H.N., Nguyen, M.H., Pham, C. 2021. Masked face recognition with convolutional neural networks and local binary patterns. *Applied Intelligence*, Springer, Online.
- [12] Ejaz, M.S., Islam, M. R. 2019. Masked Face Recognition Using Convolutional Neural Network. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 24-25 December, Dhaka, Bangladesh, 1-6.
- [13] Jeevan, G., Zacharias, G.C., Nair, M.S., Rajan, J. 2022. An Empirical Study of The Impact of Masks on Face Recognition. *Pattern Recognition*, Volume 122(2022) 108308.
- [14] Luo, X., He, X., Qing, L., Chen, X., Liu, L., Xu, Y. 2020. EyesGAN: Synthesize Human Face from Human Eyes. *Neurocomputing*, 404(2020), 213-226.
- [15] Smalakar, S., Chiracharit, W., Chamnongthai, K., Charoenpong, T. 2021. Masked Face Recognition Using Principal Component Analysis and Deep Learning. 2021 18th International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 19-22 May, Chiang, Thailand, 785-788.
- [16] Anwar, A., Raychowdhury, A. 2020. Masked Face Recognition for Secure Authentication. arXiv preprint arXiv:2008.11104.
- [17] Jang, Y., Gunes, H., Patras, I. 2019. Registration-Free Face-ssd: Single Shot Analysis of Smiles, Facial Attributes and Affect in The Wild. *Computer Vision and Image Understanding*, 182, 17-29.
- [18] Menon, A. 2019. Towards Data Science. <https://towardsdatascience.com/face-detection-in-2-minutes-using-opencv-python-90f89d7c0f81> (Erişim Tarihi: 22.04.2021)
- [19] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. 2015. Going Deeper with Convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7-12 June, Boston, MA, 1-9.
- [20] Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ICML*, arXiv: 1502.03167v3.
- [21] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. 2016. Rethinking The Inception Architecture for Computer Vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 27-30 June, Las Vegas, NV, USA, 2818-2826.
- [22] He, K., Zhang, X., Ren, S., Sun, J. 2016. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 27-30 June, Las Vegas, NV, USA, 770-778.
- [23] Yi, D., Lei, Z., Liao, S., Li, S. Z. 2014. Learning Face Representation From Scratch. arXiv:1411.7923.
- [24] Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E. 2008. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments, Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, October 2008, Marseille, France. *finria-00321923*.