



## COVID-19'un Yayılmasını Önlemek için CNN ile Yüz Maskesi Tespiti

### Face Mask Detection using CNN for Preventing the Spread of COVID-19

<sup>1</sup>Alihan DEĞİRMENCİLER , <sup>2</sup>Berat CENGİZ , <sup>3</sup>Kadir İLERİ 

<sup>1,2</sup>Karabük Üniversitesi, Mühendislik Fakültesi, Karabük, Türkiye

<sup>3</sup>Karabük Üniversitesi, Yenice Meslek Yüksekokulu, Karabük, Türkiye

<sup>1</sup>alihandegirmenciler@gmail.com, <sup>2</sup>brt\_cngz78@icloud.com,

<sup>3</sup>kadirileri@karabuk.edu.tr

Araştırma Makalesi/Research Article

#### ARTICLE INFO

##### Article history

Received : 27 August 2022

Accepted : 4 October 2022

##### Keywords:

Deep Learning, Image Processing, Face Mask Detection, CNN

#### ABSTRACT

Coronaviruses were discovered in the 1960s. This virus, which is transmitted from animals to animals, has undergone various variations until today. It came to the epidemic situation when it was first detected in a human in 2019 in Wuhan, China. There is a need to wear a mask to protect against this virus, which spreads faster in indoor areas. For this reason, wearing masks in public has become mandatory in many countries. The control of this situation has become a necessity for the countries. In this study, it was determined whether people wear masks or not using artificial intelligence and image processing techniques. The developed application detects multiple people with or without mask by using Convolutional Neural Network (CNN) in a video stream.

© 2022 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

#### MAKALE BİLGİSİ

##### Makale Tarihleri

Gönderim : 27 Ağustos 2022

Kabul : 4 Ekim 2022

##### Anahtar Kelimeler:

Derin Öğrenme, Görüntü İşleme, Yüz Maskesi Tanıma, CNN

#### ÖZET

Koronavirüsler 1960'lı yıllarda keşfedilmiştir. Hayvanlardan hayvanlara bulaşan bu virüs günümüze gelene kadar çeşitli varyasyonlara uğramıştır. Salgın durumuna ilk olarak, Çin'in Vuhan kentinde 2019 yılında bir insanda tespit edilmesiyle gelmiştir. Kapalı ortamlarda daha hızlı yayılan bu virüsten korunmak için maske takma ihtiyacı ortaya çıkmıştır. Bu nedenle, toplu ortamlarda maske takmak birçok ülkede zorunlu hale gelmiştir. Bu durumun denetimi ülkeler için gereklilik halini almıştır. Bu çalışmada, insanların maske takıp takmadığı yapay zeka ve görüntü işleme teknikleri kullanılarak tespit edilmiştir. Geliştirilen uygulama, Evrişimli Sinir Ağı (CNN: Convolutional Neural Network) kullanarak bir video görüntüsünde birçok kişiyi aynı anda maskeli veya maskesiz olarak tespit etmektedir.

© 2022 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

ORCID ID: <sup>1</sup>0000-0002-1514-1514

<sup>2</sup>0000-0002-6906-8789

<sup>3</sup>0000-0002-5041-6165

## 1. GİRİŞ

Koronavirüsünün dünya genelinde yayılarak ölümlere neden olması sonucunda Dünya Sağlık Örgütü (DSÖ) tarafından pandemi ilan edilmiştir [1,2]. Birçok ülkede özellikle kapalı alanlarda maske takma zorunluluğu getirilmiştir. İnsanların maske takıp takmadığını kontrol etmek büyük bir yük oluşturmaktadır. Bu durumun bir bilgisayar yazılım yardımıyla çözülmesi iş gücü bakımından büyük bir kolaylık sağlayacaktır.

Bir yüz maskesini tespit etmek için, nesne algılama algoritmaları kullanılabilir. Bu algoritmalarından biri olan CNN, uzamsal öznitelik çıkarma yeteneği ve daha az hesaplama maliyeti nedeniyle örüntü tanıma görevlerinde önemli bir rol oynar [2]. Çok kullanışlı uygulamalarından biri, görüntülerin ikili sınıflandırmasıdır. Qin ve Li yüz maskesi takma şeklini belirlemek için bir çalışma yapmışlardır [4]. Yüz maskesi takma şeklini, doğru maske takma, yanlış maske takma ve maske yok olmak üzere üç sınıfa ayırmışlardır. Geliştirdikleri algoritma görüntüde önce yüzü tespit eder ve sonra sınıflandırmak için SRCNet kullanır. Amit ve arkadaşları hem maskeli hem de maskesiz yüzleri algılayabilen ve CCTV kameralarla uyumlu iki aşamalı CNN mimarisi kullanan bir sistem geliştirmişlerdir [5]. Bu sistemin amacı, güvenlik ihlallerinin izlenmesine, yüz maskesi kullanımının teşvik edilmesine ve güvenli bir çalışma ortamının yaratılmasına yardımcı olmaktır. Ejaz ve arkadaşları maske tespit etme sistemi geliştirmişlerdir [6]. Temel Bileşen Analizi (PCA) uygulayarak yüzleri maskeli ve maskesiz olmak üzere iki sınıfa ayırmışlardır. Eğer yüz maskeli ise tespit etme başarısı %70'lere düşmüştür. Susanto ve arkadaşları her türlü yüz maskesini tespit edebilen bir yüz maskesi dedektörünü geliştirmişlerdir [7]. Yüz maskesini tespit etmek için YOLO V4 derin öğrenme algoritmasını seçmişlerdir. Deneysel sonuçlar gerçek zamanlı uygulamada yapılmış ve cihaz Politeknik Negeri Batam üniversitesine kurulmuştur. Geliştirilen bu cihaz, yüz maskesi takan veya takmayan kişileri çeşitli pozisyonlarda hareket etseler bile doğru bir şekilde tespit etmeyi başarmıştır. Nieto-Rodríguez ve arkadaşları sağlık personelinin tıbbi maske takıp takmadığını tespit eden bir çalışma gerçekleştirmişlerdir [8]. Bu çalışmanın ana hedefi sağlık personelinin hastanede uygun bir maske takmadığı durumda uyarı oluşturmaktır.

Bu çalışmada, insanların kapalı alanlara girişte maske takma kuralına uyup uymadığını kontrol etmeyi amaçlayan CNN ile maske tanıma yazılımı geliştirilmiştir. Geliştirilen maske tanıma yazılımı için öncelikle veri seti elde edilmiştir. Bu veri setinde maskeli ve maskesiz olmak üzere iki farklı havuz oluşturulmuş ve bu havuzlardaki veriler yapay zekanın eğitim aşamasında kullanılmıştır. Bu aşamada Derin Öğrenme'den (Deep Learning) yararlanılmıştır. Derin öğrenme, bir makine öğrenmesi yöntemidir. Verilen bir veri kümesi ile çıktıları tahmin edecek bir yapay zekayı eğitmeye olanak sağlar. Yapay zekayı eğitmek için hem denetimli hem de denetimsiz öğrenme kullanılabilir.

Yapay sinir ağları giriş katmanı, ara katman ve çıktı katmanı olmak üzere üç katmandan oluşur. Derin öğrenmedeki "Derin" ifadesi birden fazla gizli katmana sahip olma anlamına gelir. Derin öğrenmedeki en zorlayıcı kısım yapay sinir ağı eğitmektir. Bunun sebepleri:

- Büyük bir veri kümesine ihtiyaç olması,
- Yüksek miktarda hesaplama gerektirmesidir [9].

Yapay zekayı eğitmek için veri kümesindeki girdiler ve çıktılar arasında karşılaştırma yapılır. Derin öğrenmeye ihtiyaç duymamızın sebebi ise oluşan verilerin çok fazla olması ve bu verilerin işlenmesinin hızlandırılması gereğidir. Gelişen teknoloji ile GPU'ların (Ekran Kartı) gücünün artması öğrenme işleminin hızlı bir şekilde yapılabilmesine olanak tanımıştır.

Sürekli olarak güncellenen ve kendini geliştiren Python, sahip olduğu kütüphaneler sayesinde görüntü işleme teknolojisinin en rahat uygulandığı kodlama dilidir. Bu sebeple, bu proje için en uygun yazılım dili olan Python programlama dili seçilmiştir. Ayrıca, görüntü işleme kütüphanesi olan OpenCV ile derin öğrenme kütüphanesi olan Keras ve Tensorflow'dan yararlanılmıştır.

Bu çalışmanın katkısı, veri seti çeşitliliği sayesinde farklı açılardan yakalanmış görüntülerdeki kişilerin maske takıp takmadıklarını tespit edebilmesi ve maskenin uygun bir şekilde takılmadığı durumu da maskesiz olarak belirleyebilmesidir.

Bu çalışma, üç bölüme ayrılmıştır. İlk bölüm giriş ve literatür taraması, ikinci bölüm materyal ve yöntem son bölüm ise sonuç kısmından oluşmaktadır.

## 2. METARYEL VE YÖNTEMLER

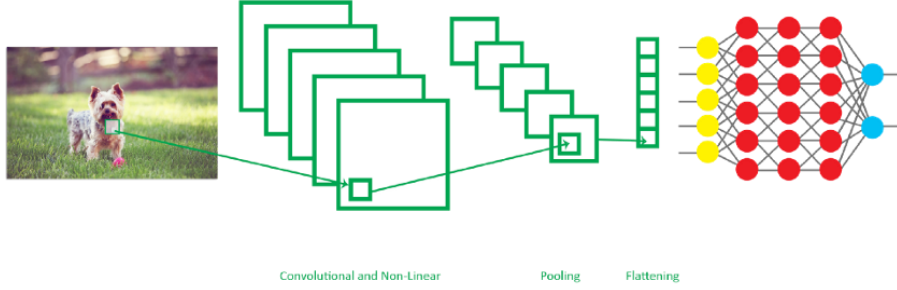
Bu bölümde Evrişimli Sinir Ağı (CNN), veri setini elde edilmesi, eğitim işlemi anlatılmıştır. Ayrıca elde edilen sonuçlar gösterilmiştir.

### 2.1. Evrişimli Sinir Ağları

Evrişimli Sinir Ağı (CNN), biyolojik bir terim gibi gözükse de bu mekanizma, nesne tanımda kullanılan çok etkili bir yöntemdir. Buradaki ana fikir bir görüntüdeki nesnelere tespit etmektir. İnsanoğlu nasıl bir kedi ile köpeği ayrı ayrı tespit edebiliyorsa, CNN yöntemi de nesnelere tespit etmek için geliştirilmiştir. Bu çalışmada maske tanımda kullanılmak üzere geliştirilen sistem, maskeyi ve yüzü birer obje olarak öğrenecek, ardından maskeli veya maskesiz olduğunu tespit edebilecektir.

CNN görüntüyü tek katmanda değil, birçok katman üzerinden inceler (Şekil 1). Bu katmanlar:

- Evrişimsel katman (Convolutional Layer): nitelikleri belirlemek için kullanılan katman,
- Doğrusalsızlık katmanı (Non-Linearity Layer): doğrusal olmayan parçanın sisteme tanıtıldığı katman,
- Ortaklama katmanı (Pooling Layer): havuz oluşturarak fazlalığı azaltır ve uygunluğu denetler,
- Düzleştirme katmanı (Flattening Layer): verilerin yapay sinir ağı girişi için hazırlandığı katman,
- Tam bağlantılı katman (Fully-Connected Layer): tam bağlı sinir ağıdır, son aşamadır [10-13].



Şekil 1. Evrişimli Sinir Ağı katmanları [10].

### 2.1.1. Evrişimsel Katman (Convolutional Layer)

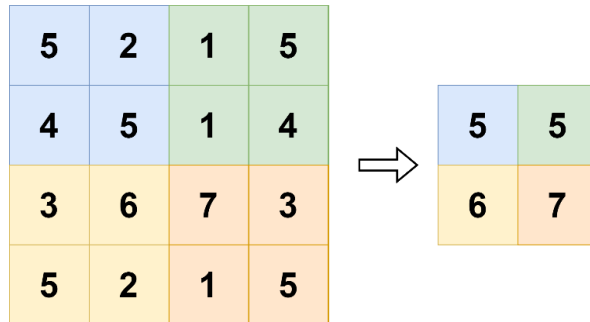
Bu katman, CNN sistemin temelini oluşturur. Fotoğrafın özelliklerini algılar, düşük ve yüksek seviyedeki özellikleri ayırt etmek için bazı filtreler uygular. Buradaki her bağlantı tıpkı insan beynindeki nöronlar gibi işler. Bir bilgi önceki katmandan sonrakine iletilir. Bu şekilde verilerin sınıflandırılması gerçekleştirilmiş olur. Her ağ verileri işleyerek özellikleri öğrenmeye başlar. Bir görüntüdeki her piksel bir giriş oluşturur. Örneğin, 1920x1080 piksel boyutundaki bir görüntü 2.073.600 ağırlığa (giriş) sahiptir. Bu durum çok işlem yükü gerektirdiğinden öğrenme işlemi çok uzun sürmektedir. Bu sorun, piksel sayısı azaltılarak çözülür [10].

### 2.1.2. Doğrusalsızlık Katmanı (Non-Linearity Layer)

Evrişimsel katmanından sonra genellikle doğrusalsızlık katmanı gelir. Görüntülerde doğrusallık problem oluşturur. Tüm katmanlar doğrusal bir fonksiyon olabildiğinden sinir ağı bunları tek bir katman olarak algılayabilir. Bu katman aktivasyon katmanı olarak ta adlandırılır. Önceleri sigmoid ve tahn gibi doğrusal olmayan fonksiyonlar kullanılmıştır, fakat en hızlı sonuç Rectifier (ReLu) fonksiyonu ile alındığından artık bu fonksiyon kullanılmaktadır [11].

### 2.1.3. Ortaklama Katmanı (Pooling Layer)

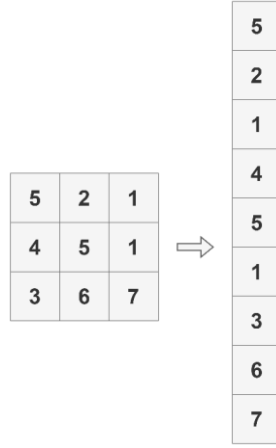
Bu katman, CNN'deki ardışık ağlar arasına eklenir. Amacı, katmanlar arasındaki hesaplama yükünü azaltmaktır. Bu sayede ağlardaki uyumsuzluk da kontrol edilmiş olur. Örneğin 4x4 boyutundaki piksellerden oluşan bir görüntüde toplam 16 ağırlık vardır. Bu görüntüye 2x2 boyutunda bir filtre (katmanlardaki en büyük sayıyı alan) uygulandığında Şekil 2'deki işlem gerçekleşir [10]. Böylece 16 ağırlık 4 ağırlığa indirgenmiş olur.



Şekil 2. Ortaklama katmanı gösterimi.

### 2.1.4. Düzleştirme Katmanı (Flattening Layer)

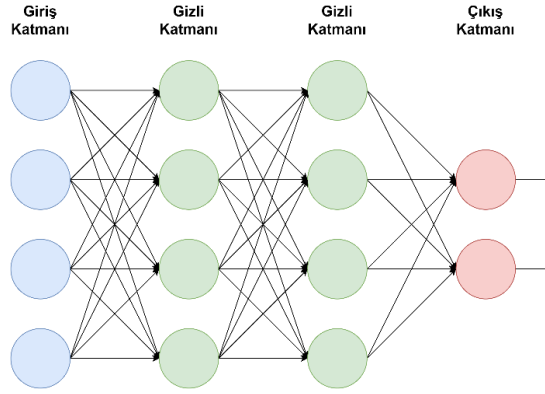
Bu katmanda ise son katman için giriş verilerini hazırlamaktadır. Genel olarak, sinir ağları giriş verilerini tek boyutlu dizi şeklinde alır (Şekil 3). Burada bulunan veriler ise bir önceki katmanlardan gelen matrislerin tek boyutlu bir diziye çevrilmiş halidir [10].



Şekil 3. Düzleştirme katmanı gösterimi.

### 2.1.5. Tam Bağlantılı Katman (Fully-Connected Layer)

CNN'in son ve en önemli katmanıdır. Öğrenme işlemi burada son bulur. Bu katmandaki düğümlerde özellikler tutulur. Evrimsel katmanından ve ortaklama katmanından geçen ve matris halinde olan veriler düz bir vektör haline getirilir. Ortaya çıkan bu vektör sinir ağına giriş olarak verilir. Yapay sinir ağının ağırlıkları ve bias (modelin yanlılık miktarı) değeri değiştirilerek öğrenme işlemi gerçekleştirilir.



Şekil 4. Tam bağlantılı katman gösterimi.

Yapay sinir ağındaki her nöron kendinden önceki katmandan alınan verilerle bilgileri aktarır ve çıktıları hesaplar (Şekil 4). Giriş değerlerine uygulanan işlev, bir ağırlık vektörü ve bir önyargı (genellikle gerçek sayılar) ile belirlenir. Öğrenme, bu ağırlıkların ve önyargıların tekrar tekrar güncellenmesiyle meydana gelir. Önyargı ve ağırlık vektörü, filtreler olarak isimlendirilir. Bu filtreler, girdinin belirli özelliklerini ifade eder. Birçok nöronun aynı filtreyi paylaşabilmesi, CNN'lerin ayırt edici bir özelliğidir. Bu durum, bellek ayak izinin azalmasını sağlar. Çünkü her alıcı alanın kendi önyargısına ve vektör ağırlığına sahip olmasının aksine, bu filtreyi paylaşan tüm alıcı alanlarda tek bir önyargı ve tek bir ağırlık vektörü kullanılır [11-14].

### 2.2. Veri Seti

Bu çalışmada, gerekli olan veri havuzu maskeli ve maskesiz görüntüler olmak üzere etiketlenerek oluşturulmuştur. Bu havuz oluşturulurken; kadın, erkek ve çocuklardan oluşan kişilerin maskeli ve maskesiz fotoğrafları kullanılmıştır. Değişik açılardan yalnızca yüz kısmını alacak şekilde fotoğraflar çekilmiştir. Bu fotoğraflar hem maskeli hem maskesiz olarak çekilmiştir. Bu fotoğraf havuzundaki fotoğrafların boyutunu azaltmak için fotoğrafların boyutları 100x60 piksel olacak şekilde boyutları küçültülmüştür (Şekil 5 ve Şekil 6).



Şekil 5. Farklı açılardan maskesiz fotoğraf örnekleri.



Şekil 6. Farklı açılardan maskeli fotoğraf örnekleri.

Veri havuzu oluştururken, kişilerin maskeyi çenesinde bırakması durumuna karşı maskenin çenede olduğu durumdaki fotoğraflarda çekilerek maskesiz olarak etiketlenmiş ve veri havuzuna eklenmiştir. Buradaki amaç, maskenin tam kapalı konumda olmadığı durumlarda da geliştirilen sistemin bu kişileri maskesiz olarak algılayabilmesini sağlamaktır. Ayrıca fotoğraflar farklı konumlarda sağa ve sola çevrilerek veri havuzu genişletilmiştir (Şekil 7).

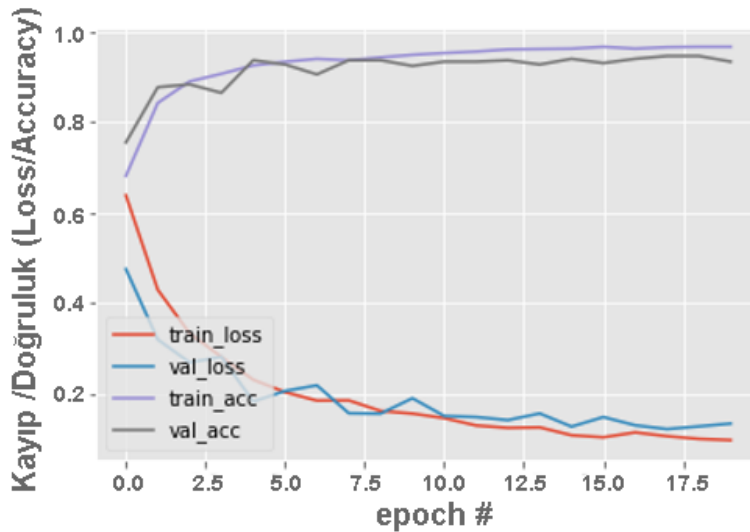


Şekil 7. Maskenin doğru takılmadığı durumlardaki fotoğraf örnekleri.

Veri havuzu oluşturulurken; maskeli 705 adet, maskesiz 998 adet fotoğraf kullanılmış ve maskenin çenede olduğu durum için de 144 adet fotoğraf kullanılmıştır. Böylece, veri havuzu toplam 1847 adet fotoğraf ile oluşturulmuştur.

### 2.3. Eğitim

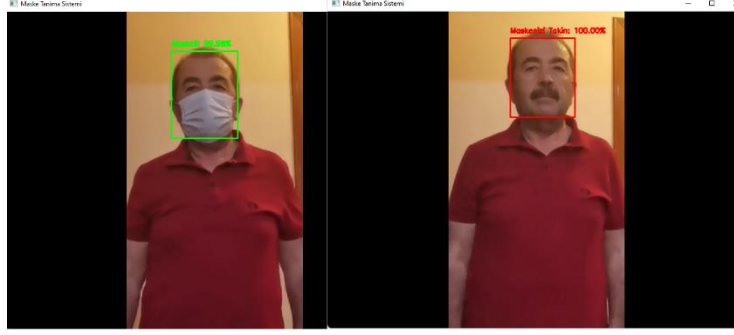
Veri seti toplandıktan sonra maskeli ve maskesiz olarak etiketlenmiştir. Daha sonra CNN kullanarak eğitim işlemi gerçekleştirilmiştir. Eğitim için, TensorFlow üzerinde çalışan Keras derin öğrenme kütüphanesinden yararlanılmıştır. Eğitim, yapay sinir ağının ağırlıklarının güncellenmesi işlemidir. Bu güncelleme işlemi her bir eğitim adımında tekrarlanır. Bu işlem, oluşturulan model için en uygun ağırlık değerleri hesaplanana kadar devam eder. Eğitim adımlarının her biri epoch olarak adlandırılır. Epoch sayısının az olması eğitim yeterli bir şekilde yapılamamasına neden olurken gereğinden fazla olması ise aşırı öğrenmeye (overfitting) neden olur. Bu iki husus istenmeyen bir durumdur ve yapay sinir ağının doğru çalışmasına engel olur. Bu yüzden optimum epoch değerinin seçilmesi gerekmektedir. Optimum epoch değeri analiz edilmiştir. Elde edilen doğruluk (accuracy) ve kayıp (loss) değerleri Şekil 8'de gösterilmiştir. Eğitim kaybının (training loss) ve doğrulama kaybının (validation loss) 20 epoch'tan sonra artmaya başladığı gözlemlenmiştir. Bu seviyeden sonra aşırı öğrenme (overfitting) gerçekleşmemesi için epoch değeri 20 olarak belirlenmiştir.



Şekil 8. Elde edilen doğruluk ve kayıp değerleri.

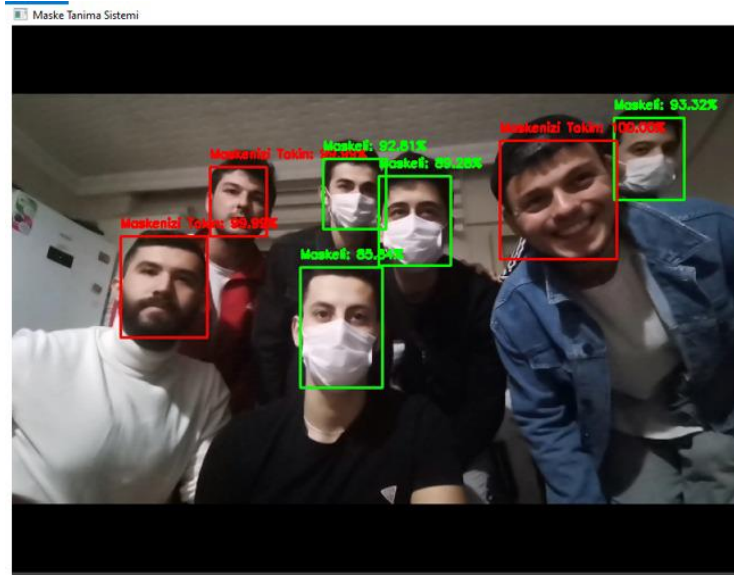
## 2.4. Geliştirilen Sistem

Veri toplama ve ardından gerçekleştirilen eğitim işleminden sonra elde edilen yapay sinir ağı bir video kameradan alınan görüntüler üzerinde insanların maskeli olup olmadıklarını tespit etmiştir (Şekil 9).



Şekil 9. Maskeli ve maskesiz durumun tespit edildiği örnekler.

Ayrıca, geliştirilen sistem, Şekil 10'da görüldüğü gibi gerçek zamanlı olarak çoklu tespit yapabilmektedir.



Şekil 10. Aynı anda birden fazla kişinin olduğu durumdaki maske tespiti.

## 3. SONUÇ

Koronavirüs nedeniyle global bir salgın oluşmuştur. Bu salgın, birçok ülkede kapalı alanlarda maske takma zorunluluğu getirilmesine sebep olmuştur. Kişilerin maske takıp takmadığının kontrolü zor bir hal almıştır. Bu durumu çözmek için yapay zeka kullanan bir sistem geliştirilmiştir. Derin öğrenme kütüphanesi Kares'i destekleyen Python programlama dili kullanarak geliştirilen bu sistem için CNN algoritması tercih edilmiştir. Maskeli ve maskesiz olmak üzere toplanan veriler ile yapay sinir ağı eğitilmiştir. Sistem, gerçek zamanlı olarak kişilerin maske takıp takmadığını tespit edebilmektedir. Maskeyi doğru bir şekilde takmayan kişileri maskesiz olarak belirleyebilmektedir. Buna ek olarak, video yayınında birden fazla kişinin bulunduğu durumda da çoklu olarak kimlerin maske takıp takmadığını yakalayabilmektedir.

Gelecekteki çalışmalarda, daha fazla veri toplanıp sistemin başarısı iyileştirilebilir. Ayrıca, maske tanımanın üzerine ek olarak ateş ölçer, kimlik kartı kontrolü gibi eklemeler yapılabilir.

## Yazar Katkıları

Yazarlar çalışmaya eşit oranlı katkı sunmuşlardır.

## Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

## KAYNAKÇA

- [1] Y.-H. Lin, C.-H. Liu, and Y.-C. Chiu, "Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries. *Brain, Behavior, and Immunity*", 2020.
- [2] H. Özbay, and A. Dalcalı, "Effects of COVID-19 on electric energy consumption in Turkey and ANN-based short-term forecasting", *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. 1, pp. 78-97, 2020.
- [3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.
- [4] B. Qin, and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19", *Sensors (Basel)*, vol. 20, no. 18, p. 5236, 2020.
- [5] A. Chavda, J. Dsouza, and S. Badgujar, "Multi-Stage CNN Architecture for Face Mask Detection", 6th International Conference for Convergence in Technology (I2CT) Pune, 2021.
- [6] M. S. Ejaz, M. R. Islam, M. Sifatullah, and A. Sarker, "Implementation of principal component analysis on masked and non-masked face recognition", 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1–5, 2019.
- [7] S. Susanto, F. A. Putra, R. Analia, and I. K. L. Nur, "The Face Mask Detection For Preventing the Spread of COVID-19 at Politeknik Negeri Batam", 2020 3rd International Conference on Applied Engineering (ICAE), Batam, Indonesia, 2020.
- [8] A. Nieto-Rodríguez, M. Mucientes, and V. M. Brea, "System for Medical Mask Detection in the Operating Room Through Facial Attributes", In: Paredes R., Cardoso J., Pardo X. (eds) *Pattern Recognition and Image Analysis*. IbPRIA 2015. *Lecture Notes in Computer Science*, vol. 9117, 2015.
- [9] N. Y. Şimsek, "Derin Öğrenme (Deep Learning) Nedir ve Nasıl Çalışır?", <https://nyilmazsimsek.medium.com/derin-ogrenme-deep-learning-nedir-ve-nasil-calisir-2d7f5850782> (Erişim Tarihi: 9 Temmuz 2022).
- [10] T. Ergin, "Convolutional Neural Network (ConvNet yada CNN) nedir, nasıl çalışır?", <https://medium.com/@tuncerergin/convolutional-neural-network-convnet-yada-cnn-nedir-nasil-calisir-97a0f5d34cad> (Erişim Tarihi: 18 Temmuz 2022).
- [11] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network", 2017 International Conference on Engineering and Technology (ICET), pp. 1-6, 2017.
- [12] A. Ajit, K. Acharya, and A. Samanta, "A Review of Convolutional Neural Networks", 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), pp. 1-5, 2020.
- [13] N. Aloysius, and M. Geetha, "A review on deep convolutional neural networks", 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 0588-0592, 2017.
- [14] Fukushima, K., "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biol. Cybernetics*, pp. 193-202, 1980.