

Received May 21, 2022, accepted June 4, 2022, date of publication June 13, 2022, date of current version June 20, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3182055

# A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System

BUSRA KOCACINAR<sup>1</sup>, BILAL TAS<sup>1</sup>, FATMA PATLAR AKBULUT<sup>1</sup>, (Member, IEEE),  
CAGATAY CATAL<sup>2</sup>, AND DEEPTI MISHRA<sup>3</sup>, (Senior Member, IEEE)

<sup>1</sup>Department of Computer Engineering, Istanbul Kültür University, 34158 Istanbul, Turkey

<sup>2</sup>Department of Computer Science and Engineering, Qatar University, Doha, Qatar

<sup>3</sup>Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjøvik, Norway

Corresponding author: Deepti Mishra (deepti.mishra@ntnu.no)

**ABSTRACT** Due to the global spread of the Covid-19 virus and its variants, new needs and problems have emerged during the pandemic that deeply affects our lives. Wearing masks as the most effective measure to prevent the spread and transmission of the virus has brought various security vulnerabilities. Today we are going through times when wearing a mask is part of our lives, thus, it is very important to identify individuals who violate this rule. Besides, this pandemic makes the traditional biometric authentication systems less effective in many cases such as facial security checks, gated community access control, and facial attendance. So far, in the area of masked face recognition, a small number of contributions have been accomplished. It is definitely imperative to enhance the recognition performance of the traditional face recognition methods on masked faces. Existing masked face recognition approaches are mostly performed based on deep learning models that require plenty of samples. Nevertheless, there are not enough image datasets containing a masked face. As such, the main objective of this study is to identify individuals who do not use masks or use them incorrectly and to verify their identity by building a masked face dataset. On this basis, a novel real-time masked detection service and face recognition mobile application was developed based on an ensemble of fine-tuned lightweight deep Convolutional Neural Networks (CNN). The proposed model achieves 90.40% validation accuracy using 12 individuals' 1849 face samples. Experiments on the five datasets built in this research demonstrate that the proposed system notably enhances the performance of masked face recognition compared to the other state-of-the-art approaches.

**INDEX TERMS** Convolutional neural networks, deep learning, fine tuning, masked face recognition, TinyML, transfer learning.

## I. INTRODUCTION

The Covid-19 virus (i.e., SARS-CoV-2) has changed the world, perhaps irreversibly and permanently. Although the treatment methods have not become widespread yet, wearing a mask, which is a relative protection method, has been made compulsory in many countries and legal action is applied in case of not wearing a mask. However, since identifying unmasked individual is both a laborious and costly task, there is a need to make it autonomous. In this case, face-mask detection based on image analysis emerges as a reliable and cheap method to protect the community health. Studies demonstrate that when an infected individual establishes close communication without wearing mask or

improper wearing of mask, the COVID-19 virus transmission is advancing rapidly. When a person sneezes or coughs, tiny droplets spread into the air and surrounding surfaces. A distance of 1.5 meters is recommended [1], as one sneeze can spread droplets up to two meters away. Contagion is not only limited to airborne transmission but also occurs via contact transmission. In addition, according to some medical professionals, when a person touches a contaminated surface in which a positive person is coughed or sneezed and then touches their mouth, eyes, or nose, this person might also be infected. As the Covid-19 information and understanding of medical professionals evolve during this pandemic, there is still more evidence for several aspects related to the Covid-19 virus. However, the science still shows that the use of face-masks reduces the spread of the virus and its variants and thereby, helps control the infection.

The associate editor coordinating the review of this manuscript and approving it for publication was Zahid Akhtar<sup>1</sup>.

In order to break the chain of infection and slow the spread of disease, advanced Artificial Intelligence techniques, particularly machine learning techniques, can be also used to endorse decisions produced by the public health system in their endeavors. With the rapid development and advancement in machine learning techniques, the dead-end of face detection systems seems to be largely addressed and that may lead the communities to the effective detection and diagnosis of this disease and also, be helpful in addressing the public problems. In the last decade, a sub-field of machine learning called deep learning algorithms have achieved many striking developments in various fields of computer vision such as image classification [2]–[4], object detection [5]–[8], segmentation [9], [10], and face recognition [11]–[15]. While traditional face recognition techniques require manual feature extraction, deep learning algorithms do not include such kinds of procedures. Deep learning algorithms can automatically extract valuable hierarchical features from training images. Especially, CNN-based algorithms provide the state-of-the-art performance in computer vision problems by applying convolution filters accompanied by various non-linear activation functions. The traditional shallow learning techniques [16]–[18] apply some filters using average or sum pooling. A common limitation of well-known face recognition algorithms is to use the face to identify individuals accurately; however, such systems assume that an image of the whole face can be taken for effective recognition. The usage of face-masks thus prevents the existing solutions to be used effectively because face masks can make the entire base around facial detection unusable.

Another consideration to construct a robust system is the size of the available training images. If there is sufficient observation in the dataset, then this can lead us to build a highly accurate model in terms of accuracy. By training on such images, the system learns to pay attention to important facial features, however, in most real-world scenarios, it is not possible to collect such kinds of large datasets or the computing power capacity might not be sufficient to process them. In addition, when such systems detect a masked face, the system cannot identify the individual. We aimed to address this security issue by ensuring that the face recognition-based system is trustworthy when presented with masked faces. The most important problem is the lack of sufficient data points for the system to be trained.

The mask detection [19]–[22] and masked face recognition [23]–[25] applications are mostly considered as separate studies in the literature. In studies where these two systems are combined, the identification of unmasked individual is the main motivation. Within the scope of this study, we proposed and validated a system to distinguish between masked, unmasked, and incorrect masked individuals by utilizing a mobile application called MadFaRe (MASkeD FACE Recognition app). In this study, a deep learning-based face recognition model has been developed. Due to the lack of face images with masks, we had to construct our dataset to train a deep network for face recognition. The fine tuned CNN-based

face recognition model has been developed and adapted to the mobile application using the service-oriented approach. The main contributions of this paper is four-fold:

- A lightweight deep learning model was built to recognize an individual's identity and detect masked faces by various mask options.
- A novel algorithm that uses only eye images to detect an individual's identity was proposed for the cases in which the calculation time is an issue due to the size of the image dataset.
- Transfer learning approach was applied to enhance the recognition performance and formulate a revolutionary paradigm associated with a face-mask identification to recognize and prevent the Covid-19 virus spread.
- A mobile application was developed to detect the improper use of face-masks in real-time to offer preventive action for the Covid-19 pandemic.

The paper is organized as follows: In Section 2, the literature of existing studies is presented; Section 3 introduces the masked face recognition system and the dataset. Section 4 presents the experimental results. Section 5 discusses a number of managerial implications and the limitations of this research. Finally, Section 6 concludes this paper and points out the future research ideas.

## II. RELATED WORK

Due to the airborne transmission of the virus, many countries around the world have introduced the face-mask as mandatory use to protect against possible infections. face-mask s basically have two main purposes of use: to prevent the transmission of viral particles circulating in the aerosol among the population and to filter volatile particles from the air. In this section, we examine the studies from the three aspects, namely, Partial face recognition, Masked face recognition, and Incorrect facemask-wearing detection.

### A. MASKED FACE RECOGNITION

The masked face recognition problem aims to match a masked face with unmasked or masked faces. This new and challenging research topic gained substantial importance during the COVID-19 pandemic when mask-wearing evolved vital to contain the spread of the virus. MFR can support the adaptability of face recognition systems in real-world scenarios, which is a compulsory component of public safety. However, the fact that the different mask types used in this process prevented the standard face recognition systems from working effectively and pushed the researchers to different solution alternatives. The habit of wearing masks during this Covid-19 pandemic has been going on for a while and unfortunately, it seems that it will continue due to new variants. As such, this has become a challenging issue to address efficiently and effectively. In the solution to this challenging problem, the fact that deep learning is a promising nominee uncovers the need for large-scale masked face datasets for training. Nevertheless, such datasets are not

readily available and adequate [26], leading researchers to adopt face masking tools to generate synthetic masked face datasets [27], [28] from existing large-scale face datasets. There are two approaches to creating masked face datasets. The first approach is to collect real-world face images with masks. However, this approach is time-consuming and costly. It is also difficult to build a large image dataset. The alternative approach is to detect the anchor points in the face images and append synthetic masks. In this approach, creating a large masked face dataset is easy due to the huge amount of open face image datasets, e.g., VGGFace and WebFace, etc. Some competitions such as International Joint Conference on Biometrics (IJCB 2021) [29] are also related to masked face recognition, which releases the corresponding dataset for research: MS-Celeb-1M [30], [31], CelebA [32], etc.

Anwar *et al.* [25] proposed a method to apply and fit masks of different colors and textures on face images based on the face landmark detector. Another group of researchers [33] studied this issue by employing a Delone Triangulation algorithm to divide the masked and unmasked face images into many small triangles, then covered each triangle of the face image with its corresponding triangle of the mask image. Additionally, some recent studies use mixed data (Masked and Unmasked face images) to train deep meshes to improve the MFR performance of the model. Organizations aiming to bring together academic and industry partners working in this field were also held. The Masked Face Recognition Competition MFR 2021 [29] has embraced these technological changes due to the widespread use of masks. It is the first competition to attract and present technical solutions that increase the accuracy of masked face recognition in real face-mask s and in a collaborative verification scenario. 18 teams participating in the competition presented solutions that aim to increase the face recognition accuracy of masked faces and deploy face recognition models in an optimized form. Out of the 18 models, 10 achieved high accuracy in improving the masked face recognition verification performance. In most solutions, architectures based on Deep Residual Networks were used to cope with the difficulty of training.

The scope of facial recognition is not limited to content taken from the front perspective. Therefore, there are also methods like HGL [34] to handle Multi-Angle Head Exposure Classification when evaluating the content presented from different perspectives. In addition to face recognition, secure authentication systems [25] have begun to be developed for masked face recognition. The main common point of these studies is the selection of deep learning techniques for implementation. In a study conducted in this context [25], tests were carried out using Multi-task Cascaded Convolutional Neural Networks (MTCNN), the distinction was made whether the individuals are masked or not, and the facial area was focused. Facial features are extracted using the Google FaceNet model, and the Support Vector Machines (SVM) algorithm is used for classification. A new facial recognition algorithm was proposed for partial occlusion using MTCNN in [35] for face detection and the LBP features of

the non-occlusion area were removed. It was reported that the achieved performance was similar to Facenet [36] for both masked and unmasked faces. In another partially occluded face recognition application [37], the occluded faces are fully recognized. In this study, a new face reconstruction algorithm is presented. This algorithm has been found to improve recognition results by up to 30%. In addition to the studies carried out with the aim of feature extraction, Openpos was used to define the mouth and extract these pixels in [21], and the CNN model was used for mask detection. Oumina *et al.* [22] combine pre-trained deep learning models such as VGG19, Xception, and MobileNetV2 with classifiers such as SVM and KNN, and then classify the extracted features under evaluation. Some studies detect masks in real-time. Suresh *et al.* [38] presented a system developed with MobileNet to detect a person who does not wear a mask and it sends the image to authorized staff members. In another study developed by Lodh *et al.* [39], a model was created by fine-tuning MobileNetv2 on a dataset consisting of masked and unmasked images in the proposed system. An accuracy of over 98% was achieved at the end of the training. After this system detects unmasked individuals, it identifies the faces. Another study uses the Multi-Task Cascaded Neural Network (MTCNN) [40] to detect the region of a face on masked face images and train the system by the LeNet algorithm used to find the accuracy differences between masked and unmasked faces. Saleh *et al.* [41] proposed a model deployed in two stages. In the first phase with 3705 images, texture and color moments feature from the face images are extracted by hybridization between these features. In the second step, the images are classified using Multi-Layer Perceptron (MLP) based on the extracted features.

## B. INCORRECT FACEMASK-WEARING DETECTION

In addition to detecting whether the face-mask has been used or not, it is also of great importance to detect when it is used incorrectly. Although most governments have forced individuals to wear face-masks, especially in public places, there are still a significant amount of people who refuse to wear or wear them incorrectly. This misuse or placement also remarkably turns down its effectiveness. This is why recent researches [27], [42], [43] have focused on detecting its placement, as well as detecting its existence. In one of the studies [28] on incorrect mask use, the authors proposed an image editing approach with the Incorrectly Masked Face Dataset (IMFD), the Correctly Masked Face Dataset (CMFD), and their combination. Tomas *et al.* [42] a method with CNN to detect incorrect usage of face-masks in real-time scenarios. Rudraraju *et al.* [27] developed an application-based face masked and incorrect use detection system from video with an accuracy of is around 90%.

When we examine the models developed to analyze whether the mask is used correctly or not, the most frequently encountered state-of-the-art models used: VGG-16, MobileNetV2, InceptionV3, ResNet-50. The obtained precision was mostly over 90% using VGG-16 and MobileNetV2.

In addition to the common base models, the recognition of the type of mask, mostly either surgical masks or N-95, has also been implemented. Such a model was developed by Qin *et al.* [43] used super-resolution and classification networks to identify the existence of a face mask. In addition to masked detection, it is capable of detecting its misuse. The study analyzed the real-life use case of the proposed model and reported that 10 images were detected in the video with 24 fps in approximately 1 second. In a recent study [44], detection was performed by considering 3 different situations. Mask versus no mask versus improper mask, mask versus no mask + improper mask, and mask versus no mask cases were analyzed over 2075 masked photographs, and 95.95%, 97.49%, and 100.0% classification accuracy were obtained, respectively. It is planned that the proposed model will be serviced and adapted to daily use. Another model [45] developed to control mask usage detected improper use of mask with very high accuracy. The classification accuracy of the model using VGG-16 was 99.81%, while the second highest model using MobileNetV2 was 99.6%. This study also aimed to differentiate whether the masks used were N-95 or surgical masks.

A recent workshop [46] was held to share face and gesture analysis in the scope of COVID-19 with international participants. Unlike MFR 2021, it focused on solving various problems of this domain on 4 datasets. One of the points where it differs from the other organization is the inclusion of improper mask usage detection. We provide a detailed comparison on different characteristics of existing studies with Table 1.

### III. METHODOLOGY

It is relatively easier to identify with the frontal view of a person because personal characteristics such as nose, eye, hair, and mouth can be deduced. However, there is a problem when it comes to recognizing individuals whose facial areas are covered by a mask in digital environments. In our study, we have proposed a two-stage deep mobile system to detect the individual who is not wearing a face mask (or wearing it improperly) and to recognize individuals from the face images. In the first stage, we gathered our faces from MaskedFace-Net [37] dataset and applied several masks (Fig. 1).

Later, we have created a three-class deep model for detecting masked, unmasked, and improper mask-wearing. Meanwhile, we have built a model that controls three conditions consisting of three types of masks. In the output of this model, information about the mask status (i.e., mask / without a mask and improper use) is obtained. In the second stage, we have constructed a face identification module that applies the following two methods: traditional identification of individuals from the whole faces and eye-based identification by getting images of the irises. The purpose of developing an eye-based recognition module is to enable mobile systems to perform fast operations with limited data.



**FIGURE 1.** Samples from the MadFare dataset that contains masked, unmasked, and incorrectly masked individuals.

#### A. DATASET AGGREGATION AND TRANSPOSITION

In this study, data were obtained from the following two datasets to detect masked faces and faces with improper worn masks: the VGGFace2 and MaskedFace-Net. VGGFace2 contains images of more than 9000 individuals of various ethnicities, professions, and ages. Over 3.3 million photographs have been recorded in total. Each person has approximately 362 photographs. MaskedFace-Net dataset contains 67,193 masked images and 69,823 improper masked images. We randomly selected several images to place a mask. We have applied face masking algorithm [25] that uses a detector of facial landmarks to specify the face tilt and key features of the face to mask faces in images. This method supports the masks such as surgical, N95, KN95, cloth, and gas. Surgical, N95, and KN95 were used for the mask detection part. Since these masks are among the most frequently used masks, we applied these three types of masks to our dataset. In our new tailored dataset (MadFare dataset), there are 151,092 images, including 50,364 with masks, 50,179 without masks, and 50,549 improperly worn masks. The number of subjects and image sizes of face recognition datasets are shown in Table 2.

To be used in face recognition training model, five distinct dataset group (VGGFace-Mini1, VGGFace-Mini2, VGGFace-Mini3, VGGFace-Mini4, MadFare) were tailored. In order to understand the effect of the volume of data, we grouped datasets from larger to small. The largest group included data over 1M, while the smallest group included 1849 images. The distribution of the datasets used for recognizing the masked face and usage type is given in Table 3.

The collected dataset aims at enabling the analyses of masked face recognition and motivates future research in image processing.

#### B. PREPROCESSING OF FACE IMAGES

In order to enhance the performance of detection of proper mask usage and the training them effectively, it is essential to preprocess the dataset. For the detection of images in actuality emphasis on image samples, a variety of dimensions is necessary. This includes rotating, scaling, translating,

**TABLE 1. Comparison with the state-of-the-art methods on masked face recognition and incorrect facemask-wearing detection.**

	Dataset	Sample Size	Method	Realtime	Mask Type	Recognition Type	Val. Acc.
Ejaz et.al. [47]	ORL face [48]	400	PCA	No	Real	Masked / Unmasked	72%, 95%
Ding et.al. [33]	CASIA-WebFace [49]	4916	Latent Part Detection	No	Augmented	Masked / Unmasked	94.34% -97.94%
Chavda et.al. [50]	Tailored datasets	7855	CNN	No	Mix	Masked / Unmasked	99.45%
Qin and Lee. [43]	Medical Masks Dataset	3835, 134, 3030	CNN	No	Real	Correct/Incorrect/Unmasked	98.70%
Rodriguez et.al. [51]	N/A	N/A	Mixture of Gaussians	Yes	Real	Masked/Unmasked	95%
Loey et.al. [52]	RMFD, SMFD, LFW	10000, 1570, 13000	Resnet50	No	Mix	Masked/Unmasked	99.49%, 100%
Tomas et al. [42]	Tailored dataset	3200 images	CNN	Yes	Real	Correct/Incorrect/Unmasked	83%
Rudraraju et al. [27]	Tailored dataset	770, 500	MobileNet	No	Augmented	Correct/Incorrect/Unmasked	90%
Cabani et al. [28]	CMFD, IMFD	137016	Haar feature cascade	No	Mix	Correct/Incorrect/Unmasked	-
Chowdary et al. [53]	SMFD	1570	InceptionV3	No	Augmented	Masked/Unmasked	99.9%
Proposed Study	VGGFace, Tailored dataset	1022811, 1849	Fine-Tuned MobileNet	Yes	Mix	MFR, Correct/Incorrect/Unmasked	90.40% 99.96%

**TABLE 2. Distributions of face recognition datasets.**

Dataset	Total Images
VGGFace-Mini1 220 subjects	1,022,811
VGGFace-Mini2 220 subjects	58,832
VGGFace-Mini3 220 subjects	58,832
VGGFace-Mini4 110 subjects	28,360
MadFare dataset 12 subjects	1849

**TABLE 3. Distributions of mask detection datasets.**

Dataset	Masked	Unmasked	Incorrectly
2 class VGGFace dataset	50,364	50,179	N/A
MaskedFace-Net dataset	50,179	N/A	69,823
3 class MadFaRe dataset	50,364	50,179	50,549

resizing images to assure that the model can recognize the masks and identity of individuals in diverse circumstances as depicted in Figure 2. To address this, we applied a series of augmentation to our face images before training the deep learning models. By preprocessing the facial images with the proposed method, the trained model can better recognize different angles and perspectives of the alike faces.

Other important processing is related to the size of the facial image that is required to be integrated prior to being used in the model. There are frequently three methods to resize images: nearest-neighbor, bilinear, and bicubic interpolation. We have used the bilinear interpolation method to resize the image due to computational advantages. The process focus is to specify the pixel values of the original image. In this context, we aligned the  $1024 \times 1024$  sized images to a  $224 \times 224$  pixels. The characteristic features of the image still exist after following the resize operation. We applied this algorithm in an automated operation to arrange the sizes. This is particularly crucial for downsizing the images for the comprehensive competence in training the model. Another element that can affect the performance of the model is the contrast ratio and luminance of the image data. In order to avoid their effects on the mask detection and face identification outputs, it is essential to regulate the contrast ratio with luminance randomly.

### C. BUILDING THE LIGHTWEIGHT DEEP MODEL FOR MASKED FACE RECOGNITION

Identification of face is carried out in two parts, uncontrolled and controlled application settings. This process is mostly used in video surveillance systems where distance of shooting, face angle, exposure, and illumination are unclear. The presence of one or more of these conditions reduces the

**FIGURE 2. Face image flipping vertically and horizontally.**

accuracy of face recognition. In addition to these, the use of mask further reduces this performance. It can also be used in many scenarios such as security checks or attendance checks in workplaces, and payments with facial-recognition. In these cases, individuals are often cooperative, typically approaching, and the face is fully visible. Thus, high-quality front face images are obtained and face recognition becomes easier. Although the mask covers most of the face, features of the upper half of the face such as eye region including eyebrows can still be used to enhance the usability of the facial recognition systems. We used two different approaches to identify the person. The first is to use the whole face and the other one is to recognize the individuals only through the eyes (Fig. 3).

Our proposed face recognition technique excels in two standpoints. One is the tailored dataset (MadFare) and another is the complete use of unexplored beneficial facial attributes. We grasped advantage of available datasets and blended them with self-generated masked faces and masked faces in real-scenes as to train the face and eye-based composite identification system. Specifically, we employed diverse attention weights to identify significant features in the masked face such as forehead, contour of face, periocular details, which productively address the problem of unbalance distribution of distinctive features. The proposed CNN-based architecture for masked face detection (Fig. 4) has the following primary components: an input layer with a pixels-wise matrix, a convolution layer, a pooling layer, a fully-connected layer, and the final output layer, which is the softmax layer.

The input layer is transformed into a three-dimensional matrix prior to feeding them into the deep network. Due to the mobile phone demands, we set the length and width of the matrix size of the image as  $224 \times 244 \times 3$  pixel values



**FIGURE 3.** Examples of a training dataset of eye detection for face recognition.

with a color channel. The convolutional layer has high importance in CNN, which attempts to analyze each small piece of input to obtain low to high level features. By moving the filter over the image and using simple matrix multiplication, features are detected. For the purpose of extracting features from simple to complex ones, different sizes of the kernels, padding, and stride values were used. In our architecture, a  $224 \times 224$  input is processed in this convolutional layer as the first layer. We set 3 as the filter size and 64 as the number of neurons. In addition to that, the pooling layer was used to lower the size of the extracted parameters and the number of calculations within the network. This operation can be considered as a transformation of a higher resolution image to a lower resolution image and it helps avoiding the overfitting problem. In this layer, we have used max-pooling to retain more texture information and average pooling to decrease the error and retain mostly background information of the image. A convolution layer with 128 neurons and a max pooling and average layer have been added. The number of neurons was increased once again, adding the third convolutional layer with 256 neurons and the max pooling layer. After proceeding through the multiple layers, a flattening layer has been added to prepare the data at the input of the most important layer, fully-connected layer. Since neural networks take the input data from a one-dimensional array, the matrices from the max pooling layer are converted to a one-dimensional array and the softmax layer turn the final results into a probability distribution matrix. Using this architecture and our three-class dataset of masked/unmasked and incorrectly worn masks, we obtained a mask detection model.

When working with small datasets, it is mandatory to deal with overfitting problem, which generally requires low complexity models. Smaller class sizes and reduced total faces loads are appropriate operative conditions. For this reason, we have reduced our class size to 12 to develop a possible solution for such small datasets. In this dataset, 1528 images were used for training and 321 images for validation. We conducted studies with three pre-trained models such as VGG16, ResNet, and MobileNet for masked face recognition. We evaluated the results obtained using various parameters such as batch size and epoch number. Size is also an important criterion as we aim to use it in the mobile application. Finally, we chose to use the pre-trained MobileNET model for our masked face recognition model (Figure 5).

We froze 23 layers of this model. We added a global average pooling, 3 dense layers, and a classification layer with 12 outputs. We obtained our masked face recognition model with our dataset, which includes masked and unmasked faces for 12 subjects' data.

An additional approach is to recognize individuals even wearing a mask through their eyes. The first is to remove the masked face area and make a recognition only over the eye area with a dataset consisting of masked and unmasked faces. First, we created a model with these two methods on a VGGFace-Mini1 and VGGFace-Mini2 datasets consisting of 220 subjects. VGGFace-Mini1 includes 1,022,811 images for masked and unmasked faces. VGGFace-Mini2 consisting of only eyes contains 58,832 images. Then, we applied Super-Resolution Generative Adversarial Networks (SRGAN) [54] to increase the resolution of our dataset consisting of eyes. Later, we decreased our data to 110 subjects and created a new model with eyes and a SRGAN application. Even using eyes only, we achieved 80.88% accuracy.

#### IV. EXPERIMENTAL RESULTS

We conducted various experiments for mask detection and masked face recognition models. In these studies, we evaluated the accuracy and loss results by decreasing the number of classes and samples to deal with a small data problem. Within the scope of the proposed system, first, we constructed the base model, and then we integrated this model into a mobile environment. In order to determine the prominent model, we carried out different experiments with the original version without fine-tuning them and with fine-tuning on masked face images, then we compared the results as discussed in the following sub-sections.

##### A. BASE MODEL SELECTION AND EVALUATION

###### 1) RESULTS OF PARTIAL FACE RECOGNITION

As a first experiment, we developed a model for recognition of individuals only from the eye region. For this model, the eye regions of 220 subjects were cropped and trained. We observed that the resolution decreased to 80.53% validation accuracy after the eye regions were cropped. In our next step, we applied SRGAN to images with reduced resolution and trained our model that reached 82.65% validation accuracy. Following that, to test the importance of the sample size, we reduced the number of subjects to 110 and we created a model using our smaller data, which consists of only eyes. As a result, we achieved an validation accuracy of 80.88% (Fig. 6).

The eye-only model offers the smallest model dimensions with very low extraction times. The results show that eye-only recognition makes it potentially convenient and practical for use on devices with low computing power.

###### 2) RESULTS OF MASKED FACE RECOGNITION

To analyze the performance of identifying faces with and without the masks, we first created a CNN-based simple

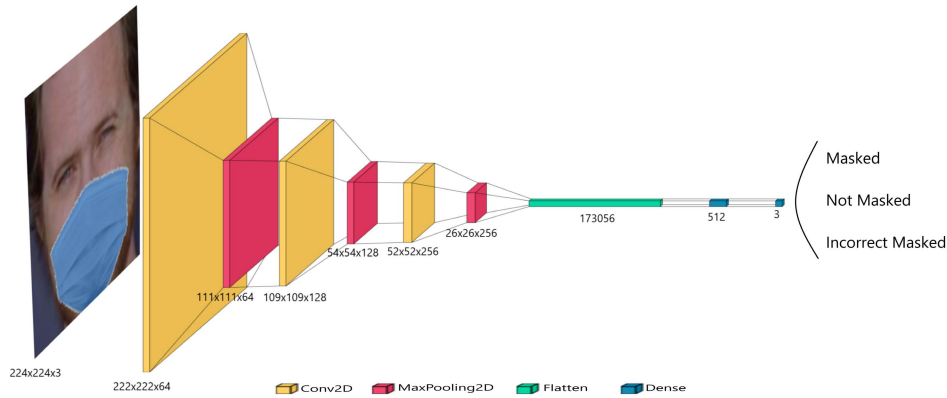


FIGURE 4. An overview of the network architecture of the proposed mask detection model.

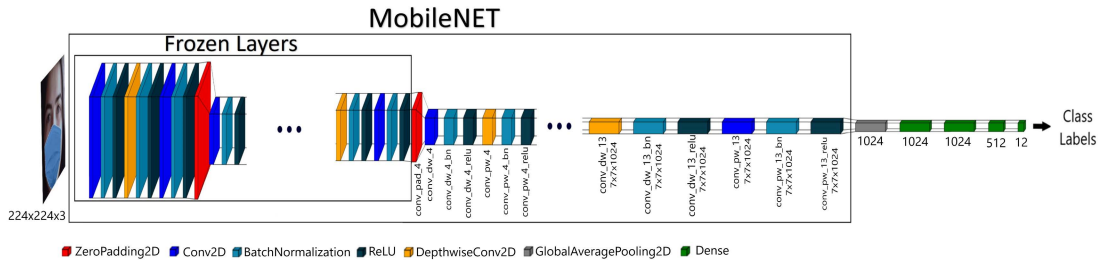


FIGURE 5. Fine-tuning on MobileNet network architecture for the masked face recognition.

model using our masked/unmasked dataset consisting of 220 subjects. The input size of this model was determined  $224 \times 224$ . Stochastic gradient descent (SGD) is selected as the optimizer. We achieved 98.92% accuracy after 10 epochs. We also worked on a reduced dataset that includes 12 subjects' images that contain masked and unmasked faces. We have reached 78.41% accuracy with this limited data and observed that reducing the input size significantly decreases the performance of the model. Nevertheless, our goal here is to construct a competitive model with a limited dataset, therefore, we continued to work on this limited dataset during the development process. The results presented in Table 4 were obtained using a variety of dataset sizes and methods.

After constructing the base model, we used pre-trained and fine-tuned pre-trained models to increase the recognition performance. Three pre-trained models were used in our study: VGG16, MobileNet, ResNet (Table 4). We used these models with variety of configuration parameters for the mobile application. As a first step, we froze the 23 layers of our pre-trained MobileNet model and fine-tuned by adding three more dense layers. We applied the L2(0.01) regularization technique and trained it for 10 epochs with 32 batch size, as such, we achieved 84.42% validation accuracy. Then, we reduced the batch size to 16 and observed a similar result with the previous case study and reached 84.43% accuracy after 10 epochs. We evaluated the results after 10 and 20 epochs and finally, we achieved 90.40% and 87.65% validation accuracy (Fig. 7), respectively.

We assessed the results with fine-tuned VGG16 architecture through three different iteration sizes, namely 100, 150, 250. As a result of this case study, we obtained the following

results for each epoch as follows: 81.40%, 82.31%, and 82.55%. To see the effect of higher iterations and batch size, we increased the batch size to 16 and then, trained the model. We achieved 77.57% validation accuracy after 250 epochs. Since the results did not improve as we expected, we continued to tune the batch size and number of iterations. After setting the batched size 32 and epoch 400, we reached 87.60% accuracy. Since we still could not get a satisfactory result, we changed the architectural model with fine-tuned ResNet and then, set the batch size to 8 with 10 iterations. As a result, we achieved 44.48% validation accuracy. Then, we created a new model by sticking to the previous configuration with the regularization technique and observed an increase over 6%. Comparison of fine-tuned architectures with a masked and unmasked dataset of 12 subjects and their results are shown in Table 5.

### 3) RESULTS OF THE INCORRECT FACEMASK-WEARING DETECTION

When developing our mask detection model, we used two datasets: MadFaRe and MaskedFace-Net. First, we conducted experiments with the MadFaRe dataset and achieved 99.93% accuracy with a two-class model produced with three kinds of masks. We repeated the same experiment with MaskedFace-Net dataset, and achieved 99.89% accuracy, which includes samples of correct or improper mask usage. As can be seen from the results, the model trained with the MadFaRe dataset offers slightly better performance than the other. As a final case study, we iterated a three-class mask detection model with MadFaRe dataset, as shown in Figure 8.

TABLE 4. Face and masked face recognition performance results on different sample size.

Method	Class	Sample	Accuracy	Validation Accuracy
Masked face recognition with VGGFace dataset	220 subjects	1,022,811 images	99.58%	98.92%
Face recognition based on the eye region	220 subjects	58,832 images	97.62%	80.53%
Face recognition based on the eye region with SRGAN	220 subjects	58,832 images	99.46%	82.65%
Face recognition based on the eye region with SRGAN	110 subjects	28,360 images	97.87%	80.88%
Masked face recognition with MadFaRe dataset	12 subjects	1849 images	94.28%	78.41%

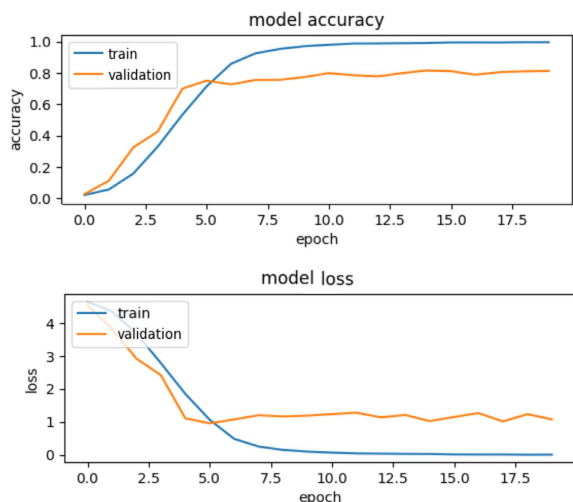


FIGURE 6. Validation accuracy and loss results: Detecting individual from eyes with 110 classes.

With the MadFaRe dataset, the estimation of only masked and unmasked individuals was realized with an accuracy of 99.84%, while it was realized with very close accuracy for 3 different classes that contain masked, unmasked, and improperly worn masked. It is an indication that the dataset and fine-tuned model proposed here can identify individuals with great success during a pandemic. We alternatively used the FMLD (Face Mask Label Dataset) [55]–[57] to test our proposed model and achieved a 0.8202 overall accuracy. We think that the reason for the decrease in Accuracy is that the patterns and motifs on the masks negatively affect the recognition of the model. Besides, the images we produce synthetically simulate only the surgical mask. In real life, the existence of different mask types such as N95, N99, w/wo ventilation, etc. creates a handicap for the model. Table 6 shows the results of different cases.

Competitive validation performance has been attained in the proposed fine-tuned solutions in comparison to the baseline model. As a result, we were able to increase the recognition accuracy of masked faces from the initial 78.41% to 90.40%.

**B. MOBILE-BASED MODEL DEPLOYMENT**

We deployed better performing prediction model to an Android-based mobile application to expedite a rapid real-time detection of mask usages. We used TensorFlow Lite to bring deep learning power into the mobile devices by running proposed model locally. It endorse hardware speed-up, limited memory usage, and leads low-latency inference efficiency to portable devices by remarkably enhancing deep

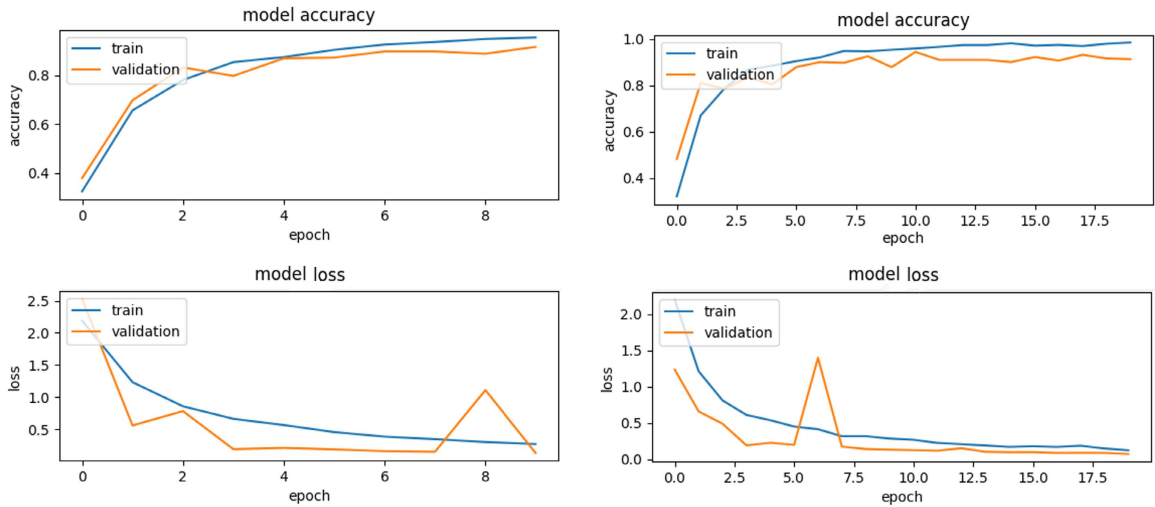
model response period. Model is deployed on the Android mobile device, face images are captured from a device integrated camera. Camera shots of a sample face image predictions are depicted in Figure 9.

The system passes through three stages after receiving the input: face detection, mask detection, and the outputs the accuracy with bounding boxes on faces. Face detection method that we preferred has features such as recognizing and finding facial features, identifying facial features, recognizing facial expressions, tracking faces in video frames, and processing video frames in real-time. All facial features are represented with 133 points in total and were used for real-time face detection in this study. After face detection, the preprocessing step is applied to resize image as 224 × 224 pixel values. Then, the boundaries of the masked face recognition system is drawn in the mobile solution as well. It is checked whether there is face data in the image taken by the system through the digital camera. If a face is detected, the position of the face is enclosed in a frame, and information about the use of the mask and identity is shown in this frame. Adhering to this system, we adapt TensorFlow’s “Object Detection Conical” example to our Android system in mobile solutions. Each face is cropped and preprocessed automatically to be fed into the model, which classifies between “masked”, “not masked” or “improper masked”. We defined two additional bitmaps for processing. The first is to rotate the input frame in portrait mode for devices whose sensor is in landscape orientation. Bitmaps are used to draw each detected face, crop its detected position, and rescale it to 224 × 224 pixels for use as the input to the face detection model. In our case, the following selections have been made: green for “Masked”, red for “Not Masked”, yellow for “incorrect Masked” and blue for “not sure” if the confidence is lower than a threshold. To create an Android application that classifies faces based on the classes in the face dataset, a text file containing the class names must exist. After the model makes a prediction on an input image, the prediction score is used to find the class label of the image. When the input value given to the model reaches a sufficient accuracy rate, it prints the identity of the person. If it provides an output value below this rate, it presents unknown. After running the application, camera can be placed in front of a face and the mobile application returns class labels that are likely to match the face.

**V. DISCUSSION**

We aimed to propose a real-time lightweight model as TinyML to evaluate the use of deep learning approaches as a





**FIGURE 7.** Validation accuracy and loss results of fine-tune MobileNet architecture with (a) 10 iteration and (b) 20 iteration to detect masked face via mobile application.

**TABLE 5.** Results of transfer learning adapted masked face recognition.

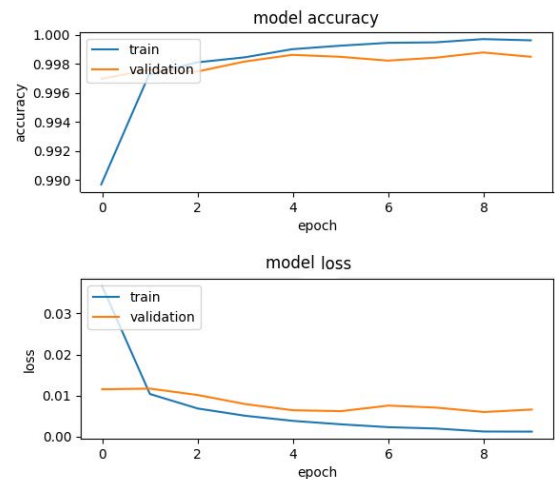
Models	Batch Size	Epoch	Accuracy	Validation Accuracy
Fine-Tuned MobileNet	32	10	94.42%	84.42%
	16	10	94.08%	84.43%
	<b>8</b>	<b>10</b>	<b>94.31%</b>	<b>90.40%</b>
	8	20	94.08%	87.65%
Fine-Tuned VGG16	8	100	90.14%	81.40%
	8	150	92.75%	82.31%
	8	250	94.03%	82.55%
	16	250	92.88%	77.57%
	32	400	92.22%	87.60%
Fine-Tuned Resnet	8	10	77.65%	44.48%
	16	20	79.78%	51.74%
Fine-Tuned Resnet(Regularization)	8	10	74.71%	51.09%

**TABLE 6.** Results of mask detection with different target classes.

Dataset	Target	Accuracy	Loss
2 class VGGFace-Masked	Is the participant masked or Not?	99.97%	0.001
2 class MaskedFace-Net	Is the participant wear the mask properly or not?	99.93%	0.004
3 class FMLD dataset	Combination of all targets:	82.02%	0.013
3 class MadFare dataset	Masked, Unmasked, and Incorrectly Masked Face	99.96%	0.001

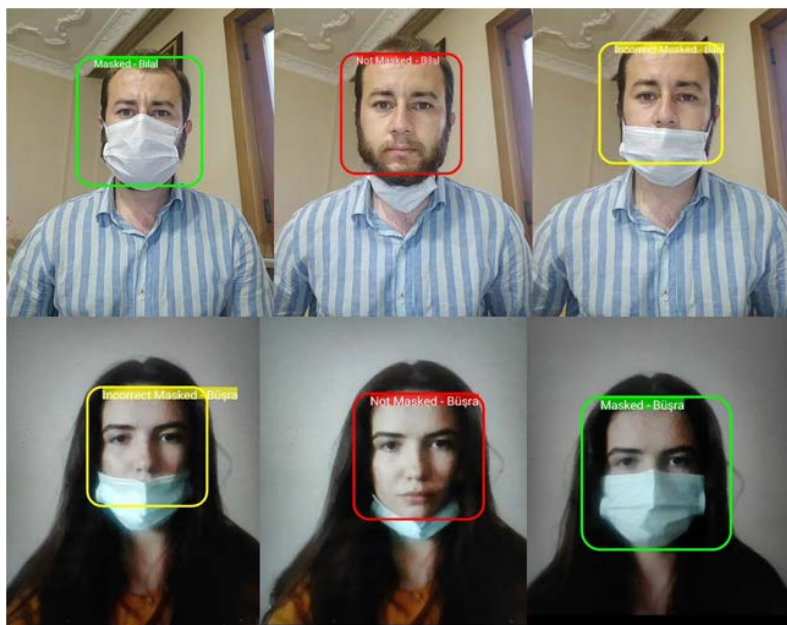
computational model and to compare the overall performance of learning on small data. Besides using only full facial images, we have also proposed a model that allows recognition from the eyes only. The proposed model has features that can run in real-time on any mobile application or in a local environment. In the proposed masked face recognition study, we observed that the performance of the model decreased as the data decreased. We used pre-trained models to increase this performance. As a result of our studies, considering the size and test results, we used the face recognition model that we created with pre-trained Mobilenet for our mobile application. The proposed mask detection model is highly precise and works without sacrificing speed in a mobile environment.

Although the proposed system is similar to the existing systems in terms of functionalities, it is a unique solution that combines these features in a single product. In addition, there are several advantages of the proposed system. with the MobileNet architecture, a low-dimensional model that can work in both mobile systems and embedded systems has been presented. The proposed model requires less computational power and resources, thereby, works faster in a

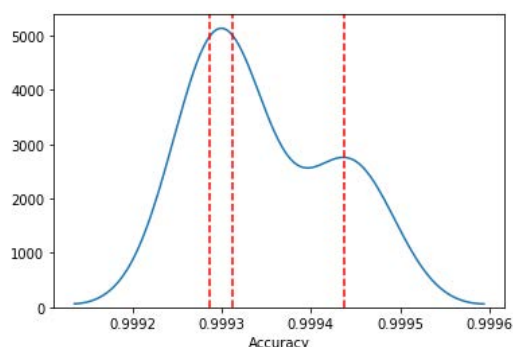


**FIGURE 8.** Validation accuracy and loss results of detecting three-class mask model (masked-unmasked-incorrectly worn mask) with MadFare dataset.

mobile system. When masked-unmasked and eye-recognition approaches of identification are compared in terms of accuracy, we observed that masked-unmasked recognition



**FIGURE 9.** Real-time masked face detection and identity identification results from the mobile application.



**FIGURE 10.** Accuracy across 10 bootstrap samples of the held out test set showing median with 95% confidence intervals.

provides better results in terms of accuracy. Applying SRGAN to increase the resolution of images on the dataset with the same number of samples had a positive effect on the accuracy. Batch size has also a noticeable effect on the accuracy rate during model training for face recognition with the masked-unmasked approach. To eliminate the effect of the distribution of skill scores, we calculated the empirical confidence intervals and plotted them in the Figure (Fig.10) below. Finally, confidence intervals shown with a 95% probability that the 99% confidence interval covers the true capability of the model are reported.

## VI. CONCLUSION

In this study, a real-time mobile masked face recognition system was presented and shown that it is an efficient method to recognize an individual. It is a general problem in such studies that the decrease in performance is due to the decrease in data. However, it is not always possible to acquire more data in the real world. On the grounds of the limited dataset, we developed an approach that could solve

this problem and proposed a model that achieves high performance even with small datasets. In our system, we fine-tuned a couple of state-of-the-art architectures such as ResNet, VGG16, and MobileNet to detect masked, unmasked, and incorrect masked usage with a limited dataset without sacrificing performance. We provided the necessary infrastructure by training the MadFaRe dataset produced with the use of three different masks in accordance with the purpose of mask recognition with our architecture, and we achieved 99.96% accuracy. Then, we limit some functionalities to adapt our model to the mobile environment and we reached 90.4% accuracy with a fine-tuned low-dimensional model MobileNet architecture. In addition to using the whole face for recognition, a model that only recognizes from the eye region has been developed to reduce the computational overhead. In order to increase the performance of the model, the resolutions of the images have been increased with SRGAN and the verification accuracy of up to 82.65% has been achieved. For this purpose, we combined both models that we developed so that we could access two fundamental functionalities to recognize masked faces and the identity of individuals at the same time to offer a real-time mobile solution that can be used in many areas where wearing a mask is mandatory for safety. This system assures a new direction towards a precise masked face recognition approach without human intervention, which is highly desirable for entries requiring biometric verification. During the global pandemic, the proposed system can be used in smartphones, surveillance cameras, and other devices.

## REFERENCES

- [1] B. Blocken, F. Malizia, T. V. Druenen, and T. Marchal, "Towards aerodynamically equivalent COVID19 1.5 m social distancing for walking and running," *Preprint*, pp. 1–12, Aug. 2020.

- [2] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. de Leon Ferreira, "Deep learning for biological image classification," *Expert Syst. Appl.*, vol. 85, pp. 114–122, Nov. 2017.
- [3] M. Pritt and G. Chern, "Satellite image classification with deep learning," in *Proc. IEEE Appl. Imag. Pattern Recognit. Workshop (AIPR)*, Oct. 2017, pp. 1–7.
- [4] Y. Hou, "Breast cancer pathological image classification based on deep learning," *J. X-Ray Sci. Technol.*, vol. 28, no. 4, pp. 727–738, Aug. 2020.
- [5] W. Budiharto, A. A. S. Gunawan, J. S. Suroso, A. Chowanda, A. Patrik, and G. Utama, "Fast object detection for quadcopter drone using deep learning," in *Proc. 3rd Int. Conf. Comput. Commun. Syst. (ICCCS)*, Apr. 2018, pp. 192–195.
- [6] F. Deng, X. Zhu, and J. Ren, "Object detection on panoramic images based on deep learning," in *Proc. 3rd Int. Conf. Control, Autom. Robot. (ICCAR)*, Apr. 2017, pp. 375–380.
- [7] A. Boukerche and Z. Hou, "Object detection using deep learning methods in traffic scenarios," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–35, Mar. 2022.
- [8] S.-P. Yong and Y.-C. Yeong, "Human object detection in forest with deep learning based on drone's vision," in *Proc. 4th Int. Conf. Comput. Inf. Sci. (ICCOINS)*, Aug. 2018, pp. 1–5.
- [9] S. Wang, D. M. Yang, R. Rong, X. Zhan, and G. Xiao, "Pathology image analysis using segmentation deep learning algorithms," *Amer. J. Pathol.*, vol. 189, no. 9, pp. 1686–1698, Sep. 2019.
- [10] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 162–169, Mar. 2019.
- [11] X. Qu, T. Wei, C. Peng, and P. Du, "A fast face recognition system based on deep learning," in *Proc. 11th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2018, pp. 289–292.
- [12] M. Arsenovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "FaceTime—Deep learning based face recognition attendance system," in *Proc. IEEE 15th Int. Symp. Intell. Syst. Informat. (SISY)*, Sep. 2017, pp. 53–58.
- [13] M. Coskun, A. Ucar, O. Yildirim, and Y. Demir, "Face recognition based on convolutional neural network," in *Proc. Int. Conf. Mod. Electr. Energy Syst. (MEES)*, Nov. 2017, pp. 376–379.
- [14] C. Karri and M. S. R. Naidu, "Deep learning algorithms for secure robot face recognition in cloud environments," in *Proc. IEEE Int. Conf. Parallel Distrib. Process. Appl., Big Data Cloud Comput., Sustain. Comput. Commun., Social Comput. Netw. (ISPA/BDCloud/SocialCom/SustainCom)*, Dec. 2020, pp. 1021–1028.
- [15] S. Manna, S. Ghildiyal, and K. Bhimani, "Face recognition from video using deep learning," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICES)*, Jun. 2020, pp. 1101–1106.
- [16] X. Wang and X. Tang, "Random sampling LDA for face recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun./Jul. 2004, pp. 1–7.
- [17] R. Gottumukkal and V. K. Asari, "An improved face recognition technique based on modular PCA approach," *Pattern Recognit. Lett.*, vol. 25, no. 4, pp. 429–436, 2004.
- [18] M. O. Faruque and M. A. M. Hasan, "Face recognition using PCA and SVM," in *Proc. 3rd Int. Conf. Anti-Counterfeiting, Secur., Identificat. Commun.*, Aug. 2009, pp. 97–101.
- [19] A. Chavda, J. Dsouza, S. Badgujar, and A. Damani, "Multi-stage CNN architecture for face mask detection," in *Proc. 6th Int. Conf. Conver. Technol. (I2CT)*, Apr. 2021, pp. 1–8.
- [20] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of coronavirus spread," *J. Biomed. Informat.*, vol. 120, Aug. 2021, Art. no. 103848.
- [21] H. Lin, R. Tse, S.-K. Tang, Y. Chen, W. Ke, and G. Pau, "Near-realtime face mask wearing recognition based on deep learning," in *Proc. IEEE 18th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2021, pp. 1–7.
- [22] A. Oumina, N. E. Makhfi, and M. Hamdi, "Control the COVID-19 pandemic: Face mask detection using transfer learning," in *Proc. IEEE 2nd Int. Conf. Electron., Control, Optim. Comput. Sci. (ICECOCS)*, Dec. 2020, pp. 1–5.
- [23] M. S. Ejaz and M. R. Islam, "Masked face recognition using convolutional neural network," in *Proc. Int. Conf. Sustain. Technol. Ind. 4.0 (STI)*, Dec. 2019, pp. 1–6.
- [24] V. Aswal, O. Tupe, S. Shaikh, and N. N. Charniya, "Single camera masked face identification," in *Proc. 19th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2020, pp. 57–60.
- [25] A. Anwar and A. Raychowdhury, "Masked face recognition for secure authentication, 2020," *arXiv:2008.11104*.
- [26] Z. Wang and T. S. Kim, "Learning to recognize masked faces by data synthesis," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIC)*, Apr. 2021, pp. 036–041.
- [27] S. R. Rudraraju, N. K. Suryadevara, and A. Negi, "Face mask detection at the fog computing gateway," in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, Sep. 2020, pp. 521–524.
- [28] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "MaskedFace-Net—A dataset of correctly/incorrectly masked face images in the context of COVID-19," *Smart Health*, vol. 19, Mar. 2021, Art. no. 100144.
- [29] F. Boutros et al., "MFR 2021: Masked face recognition competition," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–10.
- [30] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "MS-Celeb-1M: A dataset and benchmark for large-scale face recognition," in *Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer*, 2016, pp. 87–102.
- [31] X. An, X. Zhu, Y. Gao, Y. Xiao, Y. Zhao, Z. Feng, L. Wu, B. Qin, M. Zhang, D. Zhang, and Y. Fu, "Partial FC: Training 10 million identities on a single machine," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1445–1449.
- [32] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3730–3738.
- [33] F. Ding, P. Peng, Y. Huang, M. Geng, and Y. Tian, "Masked face recognition with latent part detection," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 2281–2289.
- [34] S. Li, X. Ning, L. Yu, L. Zhang, X. Dong, Y. Shi, and W. He, "Multi-angle head pose classification when wearing the mask for face recognition under the COVID-19 coronavirus epidemic," in *Proc. Int. Conf. High Perform. Big Data Intell. Syst.*, May 2020, pp. 1–5.
- [35] C. Qi and L. Yang, "Face recognition in the scene of wearing a mask," in *Proc. Int. Conf. Advance Ambient Comput. Intell. (ICAACI)*, Sep. 2020, pp. 77–80.
- [36] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
- [37] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "MaskedFace-Net—A dataset of correctly/incorrectly masked face images in the context of COVID-19," *Smart Health*, vol. 19, Mar. 2021, Art. no. 100144.
- [38] K. Suresh, M. Palangappa, and S. Bhuvan, "Face mask detection by using optimistic convolutional neural network," in *Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT)*, Jan. 2021, pp. 1084–1089.
- [39] A. Lodh, U. Saxena, A. Khan, A. Motwani, L. Shakkeera, and V. Y. Sharmasth, "Prototype for integration of face mask detection and person identification model—COVID-19," in *Proc. 4th Int. Conf. Electron., Commun. Aersp. Technol. (ICECA)*, Nov. 2020, pp. 1361–1367.
- [40] M. H. Rusli, N. N. A. Sjarif, S. S. Yuhani, S. Kok, and M. S. Kadir, "Evaluating the masked and unmasked face with LeNet algorithm," in *Proc. IEEE 17th Int. Colloq. Signal Process. Appl. (CSPA)*, Mar. 2021, pp. 171–176.
- [41] N. A. A. Saleh, H. M. Ertunc, R. A. A. Saleh, and M. A. Rassam, "A simple mask detection model based on a multi-layer perception neural network," in *Proc. Int. Conf. Technol., Sci. Admin. (ICTSA)*, Mar. 2021, pp. 1–5.
- [42] J. Tomás, A. Rego, S. Viciano-Tudela, and J. Lloret, "Incorrect facemask-wearing detection using convolutional neural networks with transfer learning," *Healthcare*, vol. 9, no. 8, p. 1050, Aug. 2021.
- [43] B. Qin and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19," *Sensors*, vol. 20, no. 18, p. 5236, Sep. 2020.
- [44] E. Aydemir, M. A. Yalcinkaya, P. D. Barua, M. Baygin, O. Faust, S. Dogan, S. Chakraborty, T. Tuncer, and U. R. Acharya, "Hybrid deep feature generation for appropriate face mask use detection," *Int. J. Environ. Res. Public Health*, vol. 19, no. 4, p. 1939, Feb. 2022.
- [45] S. Hussain, Y. Yu, M. Ayoub, A. Khan, R. Rehman, J. A. Wahid, and W. Hou, "IoT and deep learning based approach for rapid screening and face mask detection for infection spread control of COVID-19," *Appl. Sci.*, vol. 11, no. 8, p. 3495, Apr. 2021.
- [46] H. K. Ekenel, F. İ. Eyiokur, M. Imran, A. Kantarcı, F. Ofli, V. Štruc, A. Waibel, O. Aran, A. Bharati, and C. Busch, "Workshop committees FG 2021," in *Proc. 16th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, 2021, p. 1.

[47] M. S. Ejaz, M. R. Islam, M. Sifatullah, and A. Sarker, "Implementation of principal component analysis on masked and non-masked face recognition," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–5.

[48] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Dec. 1994, pp. 138–142.

[49] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," 2014, *arXiv:1411.7923*.

[50] A. Chavda, J. Dsouza, S. Badgajar, and A. Damani, "Multi-stage CNN architecture for face mask detection," in *Proc. 6th Int. Conf. Conver. Technol. (I2CT)*, Apr. 2021, pp. 1–8.

[51] A. Nieto-Rodriguez, M. Mucientes, and V. M. Brea, "System for medical mask detection in the operating room through facial attributes," in *Proc. Iberian Conf. Pattern Recognit. Image Anal.* Cham, Switzerland: Springer, 2015, pp. 138–145.

[52] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement*, vol. 167, Jan. 2021, Art. no. 108288.

[53] G. J. Chowdhary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Face mask detection using transfer learning of Inceptionv3," in *Proc. Int. Conf. Big Data Anal.* Cham, Switzerland: Springer, 2020, pp. 81–90.

[54] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.

[55] B. Batagelj, P. Peer, V. Štruc, and S. Dobrišek, "How to correctly detect face-masks for COVID-19 from visual information?" *Appl. Sci.*, vol. 11, no. 5, p. 2070, Feb. 2021.

[56] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with LLE-CNNs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2682–2690.

[57] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A face detection benchmark," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5525–5533.



**FATMA PATLAR AKBULUT** (Member, IEEE) received the B.S. and M.S. degrees in computer engineering from Istanbul Kültür University (IKU), and the Ph.D. degree in biomedical engineering from Istanbul University, in 2017. Her Ph.D. work focused on developing machine learning enabled wearable systems for long-term cardiovascular disease monitoring. She is currently an Assistant Professor with the Computer Engineering Department, IKU. Prior to IKU, she worked as a Postdoctoral Researcher at the Advanced Self-Powered Systems of Integrated Sensors and Technologies (ASSIST) Center, Computer Science Department, North Carolina State University (NCSSU), from 2017 to 2019. Her research interests include biomedical signal processing, affective computing, data analytics, and wearable systems for healthcare.



**CAGATAY CATAL** received the B.S. and M.Sc. degrees in computer engineering from Istanbul Technical University and the Ph.D. degree in computer engineering from Yıldız Technical University, Istanbul. He was a Full Professor at the Department of Computer Engineering, Bahcesehir University, Istanbul, from 2020 to 2021. He has been working as an Associate Professor at the Department of Computer Science & Engineering, Qatar University, since January 2021. He worked two years as a full-time Faculty Member at Wageningen University & Research (WUR)/Information Technology Chair Group in The Netherlands before joining Bahcesehir University. Before WUR, he worked six years at the Department of Computer Engineering, Istanbul Kültür University, as an Associate Professor and the Head of the Department. He received the Associate Professor title from the Inter-University Council of Turkey, in April 2014. Before joining the university, he worked eight years at the Scientific and Technological Research Council of Turkey (TUBITAK), Information Technologies Institute, as a Senior Researcher and a Project Manager. His research interests include deep learning, machine learning, natural language processing, software engineering, software testing, and software architecture. He is an External Reviewer of the EUROSTARS Program funded by the European Union, the Research Council of Canada, and the Research Council of Turkey. His recent publications can be accessed from the following page: <http://publicationslist.org/cagataycatal>.



**BUSRA KOCACINAR** received the B.S. degree in computer engineering from Istanbul Kültür University, where she is currently pursuing the M.S. degree with the Computer Engineering Department. She also works as a Research Assistant at the Computer Engineering Department, Istanbul Kültür University. Her research interests include artificial intelligence, big data, and machine/deep learning.



**BILAL TAS** received the B.S. degree in computer engineering from Istanbul Kültür University, in 2021. He continued to work actively as a Software Developer at Vestel while he was still a student and currently working as a Software Engineer in a defense industry company. His research interests include software engineering and machine learning applications.



**DEEPTI MISHRA** (Senior Member, IEEE) has been working as an Associate Professor with the Department of Computer Science (IDI), Norwegian University of Science and Technology (NTNU), since 2016. She is currently the Head of the Software, Data, and Digital Ecosystem (SDDE) Research Group and the Educational Technology Laboratory at the Department of Computer Science, NTNU. She has an extensive international experience and earlier worked at Monash University Malaysia; Atilim University, Turkey; and various institutions at India. Her research interests include empirical software engineering, robotics, human–robot interaction, artificial intelligence, sustainability, and computing education.

...