Research Article
Received: 17.04.2023
Accepted: 27.03.2024
Published: 29.03.2024

Journal of Disaster and Risk
7(1), 2024, (71-85)
e-ISSN: 2636-8390
DOI: 10.35341/afet.1284220

Determining the Reliability of Personal Masks with Convolutional Neural Networks

Özgür Boran Ak[1], Ertan Kuruöz[2], Ayça Ak[3]

Abstract

During the COVID-19 pandemic, which is a worldwide disaster, it has been proven that one of the most important methods to struggle the transmission of such diseases is the use of face masks. Due to this pandemic, the use of masks has become mandatory in Turkey and in many other countries. Since some surgical masks do not comply with the standards, their protective properties are low. The aim of this study is to determine the reliability of personal masks with Convolutional Neural Networks (CNNs). For this purpose, first, a mask data set consisting of 2424 images was created. Subsequently, deep learning and convolutional neural networks were employed to differentiate between meltblown surgical masks and non-meltblown surgical masks without protective features. The masks under investigation in this study are divided into 5 classes: fabric mask, meltblown surgical mask, meltblown surgical mask, respiratory protective mask and valve mask. Classification of these mask images was carried out using various models, including 4-Layer CNN, 8-Layer CNN, ResNet-50, DenseNet-121, EfficientNet-B3, VGG-16, MobileNet, NasNetMobile, and Xception. The highest accuracy, 98%, was achieved with the Xception network.

Keywords: Artificial Intelligence, Convolutional Neural Networks, Image classification, Personal Mask

1. INTRODUCTION

It is reported that COVID-19/coronavirus disease, which occurs with acute respiratory symptoms (SARS-CoV2) infects more than 250 million people global and reasons more than 5 million deaths. Strengthening health systems, maintaining social distancing and improving surveillance are the keys to controlling the COVID-19 pandemic. World Health Organization (WHO) states that the main measures that can be taken to stop the more spread of this fatal virus are to wear a face mask, along with maintaining proper social distance. Research on measures to pandemic the COVID-19 pandemic reveals that a suitable face mask that covers the nose and mouth reduces the risk of spreading the coronavirus by more than 90%. Public health and government agencies recommend face masks as basic measures to keep the public safe. Most countries have made it obligatory to wear masks both indoors and outdoors.

When the studies on masks are examined, there are studies that distinguish those who wear masks and those who do not, or that the mask is not worn correctly. So et al. detected who uses and does not use face masks by using EfficientNet and YOLOv3 algorithms; classified who uses face masks as qualified mask users and unqualified mask users (Su et al., 2022). Köklü et al. (2022)

[1] Prof. Dr. Şaban Teoman Duralı Science and Art Center, Kartal, İstanbul, Türkiye
e-mail: ozgurboranak@gmail.com, ORCID NO: 0000-0002-2523-7640
[2] Prof. Dr. Şaban Teoman Duralı Science and Art Center, Kartal, İstanbul, Türkiye
e-mail: ertankuruoz@hotmail.com, ORCID NO: 0000-0002-9628-7418
[3] Doç.Dr. Marmara University Vocational School of Technical Sciences, 34865 Kartal, İstanbul, Türkiye
Corresponding author e-mail: aycaak@marmara.edu.tr, ORCID NO: 0000-0002-3429-4962

To cite this article
Ak, Ö. B., Kuruöz, E., Ak, A., (2024). Determining the Reliability of Personal Masks with Convolutional Neural Networks. *Journal of Disaster and Risk,* 7(1), 71-85.

classified as who wears masks correctly, whose masks are under nose, whose masks are under chin, who does not use mask. Their dataset consists of 2000 images. The classification layers of AlexNet and VGG16 models were removed and replaced by Long-Short Term Memory and Bidirectional Long-Short Term Memory architectures. They were achieved 95.67% accuracy. Kansal et al. (2021) compared the performance of different machine learning algorithms detecting face masks. Tomas et al. (2021) aimed to develop an intelligent method for automatically detecting incorrectly wearing face masks. In their study, they used transfer-learning CNN to detect not only whether a mask was used, but also added errors that could contribute to the spread of the virus. In this study, the places where the masks were worn incorrectly were determined by labeling the pictures. They used 3200 photographs, which they obtained with public participation through an application they developed, as a database. They achieved the highest accuracy of 83.4% with the VGG16 model. Snyder and Husari (2021) have presented an end-to-end approach to deep learning-based face mask detection for poor quality images from difficult distances, angles, and available lighting quality. In their study, the first component uses deep residual learning (ResNet-50) to detect the presence of human subjects in the videos, and the second component uses Multitask CNN to identify and take out human faces from these videos. In the last component, CNN is used to detect masked and unmasked people. Sethi et al. (2021) detected person that does not wear facemasks. They achieved a 98.2% accuracy in discerning the presence of a mask using the Resnet50 model. Introducing a novel object detection approach that amalgamates single-stage and two-stage detectors, they successfully identified objects in real-time video streams. Additionally, they annotated their work and employed 14 million images from the ImageNet dataset.

Teboulbi et al. (2021) detected masks and social distance using pre-trained models such as ResNet and MobileNet classifiers. They utilized a dataset comprising 3835 images and employed MobileNet, ResNet Classifier, and VGG models. With these models, they achieved a remarkable 99% accuracy in determining both social distancing compliance and the presence of masks. Chavda et al. (2021) detected people that does not wear face masks. Detection of both masked and unmasked faces was made using a two-stage CNN and their method is compatible with CCTV cameras. They determined whether a mask was worn with around 99% accuracy using NASNETMobile, DenseNet121 and MobileNetV2 models with labelling. Song et al. (2022) detected people wearing face masks, the position of the mask, classified mask type as homemade, surgical and n95 mask, recognized the person's identity. In their study, Facial Recognition Pipeline with FaceNet, AlexNet, CNN and VGG16 were used to classify features. In addition, they performed classification of N95s, cloth, and surgical masks. Their dataset comprised 67,193 images of correctly worn masks and 66,900 images of incorrectly worn masks. Across various models, they attained an average accuracy of 97%. Rokhana et al. (2021) performed a multi-class image classification study to determine the correct use of the face mask based on the MobileNetV2 architecture. O'Kelly et al. (2021) compared the face fit of N95, KN95, surgical and fabric face masks using Respirator Fit Tester. Wakarekar and Gurav's (2022) research presented a face mask classification algorithm that combines Efficient-Yolov3, deep learning and the viola Jones method to determine if the mask exists. Goyal et al. (2022) presented a face mask detection model that classifies images as masked and unmasked for static and real-time videos. Their model was trained using the Kaggle dataset of 4000 images and achieved an accuracy rate of 98%.

Kayalı et al. (2021) classified people wearing masks correctly, wearing masks wrongly and are not wearing a mask. Kaur et al. (2022) used some machine learning tools as Keras, TensorFlow, Scikit-Learn and to recognize the face in the video or image. They determined whether or not it has a mask on it. Asif et al. (2021) proposed a system with two components. The first component is for the face detection and the other is to identify the mask area. They used MobileNetV2 structure and deep transfer learning model to identify the mask area. In the study of Sharma et al. (2022) a solution system is designed over Internet of Things (IoT) for face mask detection and classification. The system used MobileNetV2, VGG-16, ResNet-50, Inception v3, and CNN models and proposed to identify persons who wear the face mask. Rajath et al. (2021) suggested a model for mask face detection established on deep learning and computer vision. Their system can be inserted with a surveillance camera to detect people with mask or not face masks. They integrate classic machine learning methods and deep learning using open computer vision, keras and tensor flow. The work of Naufal et al. (2021) proposes to compare the classification methods of classical machine learning to classify faces with mask and without mask pictures. They used CNN, support vector machine (SVM) and k-nearest neighbours (KNN) methods for detection of facemasks. They achieved the highest accuracy (96.83%) with CNN. Table 1 summarizes some studies and compares each study according to the problem it focuses on, the models it uses, accuracy, and advantages/disadvantages.

Table 1. Literature Survey (Ciuffreda, 2021).

Reference	Year	Type of Detection	Used Models	Best Accuracy	Advantages/Disadvantages
Tomas et al.	2021	Incorrect facemask wearing	Xception InceptionV3 MobileNet, ResNet50 NASNetLarge VGG16	83.4%	good performance in determining whether a mask is worn/poor performance in determining mask class
Tebouldi et al.	2021	Whether a mask is worn or not social distance	MobileNet, ResNet VGG	99%	Good Performance/ requires additional hardware
Song et al.	2022	Mask Detection, Mask Type Classification, Mask Position Classification, Identity Recognition	FaceNet, AlexNet, CNN VGG16	97%	Offering a comprehensive system/ Requires a complex infrastructure
Sethi et al.	2021	with/without mask	AlexNet MobileNet ResNet50	98.2%	requires less memory/ too many images
Naufal et al.	2021	with/without mask	KNN SVM CNN	96.83%	Trying methods other than CNN/ Trying a single CNN model
Köklü et al.	2022	Whether a mask is worn or not social distance	AlexNet VGG16	95.67%	Using LSTM and BILSTM/ Decreased accuracy in some classifications

Since the days when personal masks became mandatory during the Covid epidemic, many studies have been carried out on distinguishing individuals who wear masks and those who do not, or whether the mask is worn correctly or not. However, there are not enough studies on whether the mask worn has sufficient protection. In contrast to existing literature, this study aimed to classify masks directly. For this purpose, a mask database consisting of photographs of fabric masks, meltblown surgical masks, and surgical masks without meltblown masks was created. Unlike other studies that determine whether individuals wear masks, this research allows for the assessment of the quality of a mask, enabling its evaluation before being worn on the face. Additionally, it determines whether the acquired mask is protective.

In the second phase of the study, various CNN models such as 4-layer CNN, 8-layer CNN, ResNet-50, DenseNet-121, EfficientNet-B3, VGG-16, MobileNet, NasNetMobile, and Xception were employed to classify mask photos using the prepared database. In contrast to many studies using different CNN architectures for classification, an additional layer has been added to the network that yielded the highest accuracy score, further enhancing the accuracy through fine-tuning. The results are presented comparatively.

The remainder of this article is organized as follows: Section 2 provides material and the methods. In section 3, the results of the analysis are explained. Section 4 concludes the study.

2. MATERIAL AND METHOD

The block diagram of the proposed method is given on Fig. 1. After the dataset consisting of masks was prepared, pre-trained and conventional network structures were trained. The method with the highest accuracy was determined by comparing the results. At this stage, after adding a fully connected layer, fine-tuning has been made and the masks have been classified with the highest accuracy.

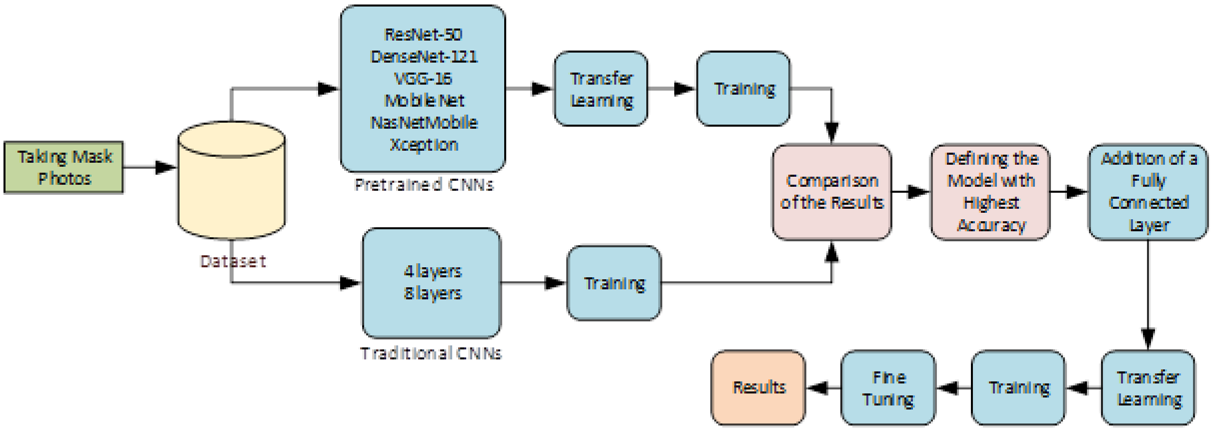
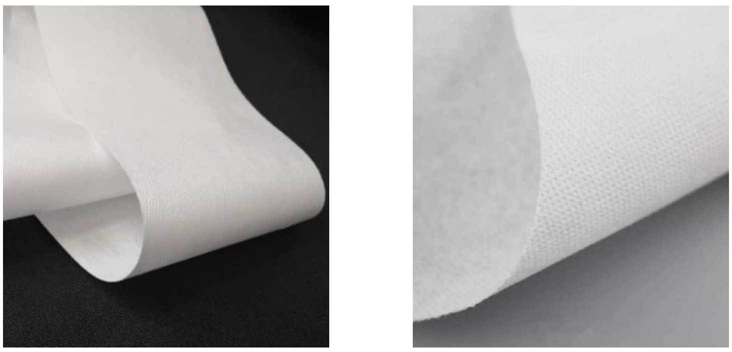


Figure 1. Proposed method.

2.1. Masks

Due to the Covid-19 pandemic, the use of masks has become mandatory in Turkey and many countries. The outer layers of surgical masks should consist of a spunbond layer that protects the mask from water and moisture, and the inner layer of a meltblown layer that has a filter feature. Figure 2 shows spunbond and meltbown fabrics (URL 1).

There is EN 14683 standard for surgical masks. According to this standard, there are two classes of surgical masks, type I and type II. Type II masks, on the other hand, have two standards, type II and type IIR, according to their splash resistance. Type I masks should have at least 95% bacterial filtration and type II masks should have at least 98% bacteria filtration. Type IIR masks, must be splash resistant at a pressure of at least 16 kPa. Performance requirements for medical face masks are given in Table 2.



a) Spunbond b) Meltbown

Figure 2. Mask fabrics.

Table 2. Medical face mask performance requirements (Ciuffreda, 2021).

Test	Type I	Type II	Type IIR
Bacterial filtration (%)	≥95	≥98	≥98
Differential pressure (Pa/cm²)	<40	<40	<60
Bounce resistance pressure (kPa)	Not required	Not required	≥16
Microbial cleaning (cfu/g)	≤30	≤30	≤30

However, some surgical masks do not comply with these standards, all three layers of these masks consist of spunbond fabric, and their protective properties are low. In addition, the protection of different types of masks is different. Cloth masks have low protection as they are made of fabric and do not contain meltblown. Respiratory protective masks have higher protection than surgical masks. There is EN 149 standard for these masks. According to this standard, there are three classes of respiratory protective masks: FFP1, FFP2, FFP3. FFP1 masks should have at least 80% particle filtration, FFP2 masks should have at least 94% and FFP3 masks should have at least 99% particle filtration. Although valved masks are respiratory protective masks, they have low protection against contamination since they only filter the air while inhaling and exhale directly.

Cloth mask, surgical mask with meltblown, surgical mask without meltblown, protective respiratory mask (N95), mask with valve are shown in Figure 3.

Figure 3. Mask types.

A total of 2424 mask photographs were taken to create the dataset. Mask photos were taken from different lights and angles. Some of the images were taken by holding the masks to the light as the texture of the mask is clearer. After the models are trained, 223 images are shot and masks in the images are coloured in computer to train the model with coloured masks. The model with highest accuracy is trained with the new images added to the dataset. Clear sky, cloudy sky, white bulb, yellow bulb light sources are used. Some images were shot at high ISO speed for noise purposes.

Images were taken at 1984x2976x3 resolution using Canon eos 200d camera. The resolutions of the images were reduced with a code prepared in Python. After the resolution of these images was reduced to 1000x1500x3, they were resized according to the input resolution of the relevant network during training. The number of photographs according to mask types is given in Table 3.

Table 3. Number of Photographs

Mask Types	Number of Photographs
Cloth Mask	529
Surgical Mask With Meltblown	482
Surgical Mask Without Meltblown	441
N95	803
Mask With Valve	410

Rotation, mirroring and translation data augmentation methods were also implemented with code written in Python.

2.2 Convolutional Neural Networks

Artificial neural networks (ANN) is an information processing technology that imitates the way the biological nervous system works, inspired by the working technique of the human brain. In an ANN, the input signal is multiplied by a weighting coefficient and sent to the sum function. A fixed bias value is added to this, independent of the input values. This weighted sum is passed through the activation function. The most used activation functions are logistic, sigmoid, ReLU (Rectified Linear Unit) and tanh or hyperbolic tangent functions.

Developed from ANN, deep learning is the latest attainment in the field of machine learning where it offers superhuman abilities in many applications such as object recognition, anomaly detection, and emotion recognition. Deep learning models have of interconnected several hidden layers. Therefore, the term "deep" learning is a specific branch of machine learning that can deal with complex patterns and objects in large datasets. CNNs are feedforward ANNs with alternating layers of convolution and subsampling and learn based on the visual perception of living things. Different CNN models are used especially in computer vision and human activity classifying problems (Sighencea, et al., 2021; Bozkurt, 2022).

A CNN consists of three types of layers (Kiranyaz, et al., 2021). These can be listed as fully connected layers, pool layers and convolution layers. CNN architecture is shaped by the combination of these layers in different numbers (O'Shea et al., 2015). The basic CNN model architecture is shown in Figure 4 (URL 2).

Convolution is an operation in which two pieces of information are intertwined, transforming one function into another. Convolutions are typically used in image processing to blur and sharpen images as well as enhance edges and emboss. For each convolution kernel k_j and input x_j, the output feature map is defined by the following Equation 1:

$$y_{i,j} = f\left(b_j + \sum_i w_j * x_i\right) \qquad (1)$$

where, b is the bias, w is the weights, f is the activation function and * symbolizes the convolutional operation.

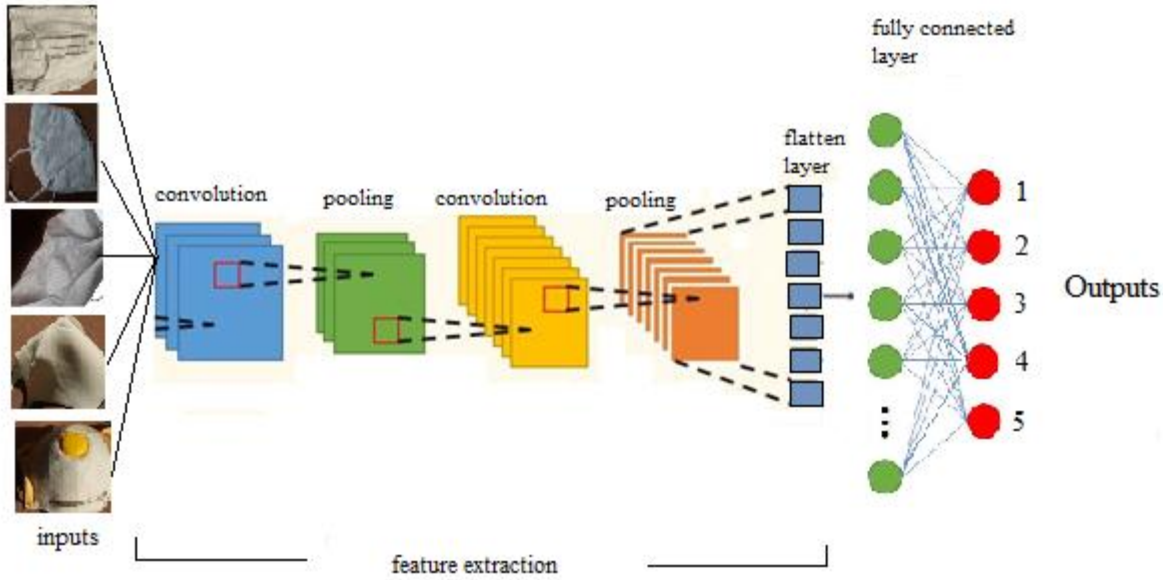


Figure 4. CNN architecture.

All neurons in this layer perceive the same feature only at different positions at the input. Activation functions are operations placed at the end or between layers and decide whether the neuron should fire or not. The selection of the activation function in both the output layer and the hidden layers is important, as the model controls the quality of learning. Softmax (Eq. 2) is the activation function used for the output layer, while ReLU (Eq. 3) is preferred for the hidden layers:

$$\textit{Softmax:}\ (x_i) = \frac{e^{-(x_i)}}{\sum_j e^{-(x_i)}} \qquad (2)$$

$$\textit{ReLU:}\ f(x) = max(0, x)\,, \quad x > 0 \qquad (3)$$

Pooling layers reduce the size of the feature maps. Thus, the number of parameters to be learned and the amount of calculations made on the network are reduced. Two common pooling methods are average pooling and maximum pooling. y_i output feature map can be written as Equation 4 for x_j input feature map:

$$y_i = \max_{rxr}(x_i) \qquad (4)$$

where r represents the pooling size (Du, et al., 2016).

2.3 Transfer Learning

Transfer learning is to train a model that has trained with a large dataset (e.g. Imagenet) instead of training a model with randomly initialized weights. It helps delaying overfitting and getting better accuracy with a small dataset. Transfer learning has two steps: First step is freezing the convolutional and batch normalization layers of the network, replacing the pretrained fully connected layers with appropriate fully connected layers for the problem and training the fully connected layers. The convolutional layers are frozen to preserve the pretrained features from large gradient updates that is caused by randomly adjusted fully connected layers. The second step is fine tuning: Fine tuning is to unfreeze the convolutional layers of the model and training the whole model with a low learning rate. Batch normalization layers are not unfrozen because they destroy the training that has been done before. A much larger model is training in this step so using a lower learning rate is important in fine-tuning.

ResNet-50: The main feature of the ResNet architecture (Fig. 5) is that the network can be trained with higher accuracy and faster as the number of layers increases. ResNet offers an easier training process by adding skip connections between layers. ResNet-50 architecture consists of convolution and identity sections. ID blocks are a standard block used in ResNet. The first component of the identity block is the convolution layer. Then, batch normalization is used. Here, RELU is used as the activation function. The second component is similar to the first except for filter size. The third component has no activation function. The convolutional block follows the same process as the identity block with an additional 3D convolution layer (Fırat and Hanbay, 2023).

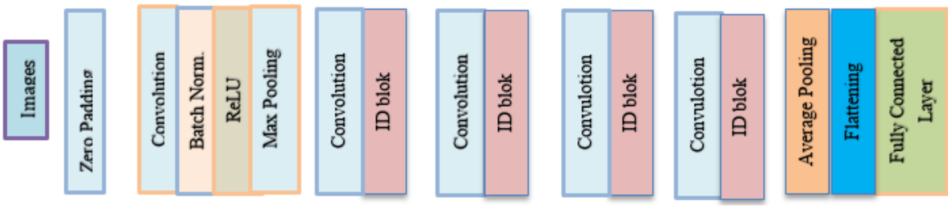


Figure 5. ResNet-50 architecture.

DenseNet-121: DenseNet is one of the modern architectures of CNN for image recognition with fewer parameters. DenseNet combines the previous layer output with a subsequent layer, along with its concatenated features. For this purpose, an additional attribute is used to combine the previous layer with future layers. DenseNet architecture (Fig. 6) aims to solve this problem by densely interconnecting all layers. Among different DenseNet (DenseNet-121, DenseNet-160, DenseNet-201), DenseNet-121 architecture was used in this study. The details of DenseNet-121 are as follows: 5 convolution and pooling layers, 3 transition layers (6,12,24), 1 Classification layer (16) and 2 dense blocks (1 × 1 and 3 × 3 convolution).

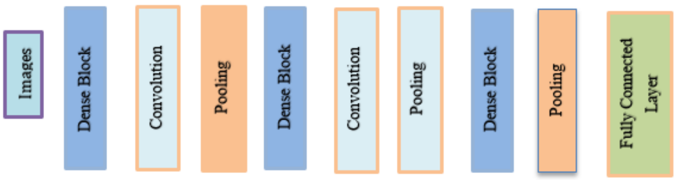


Figure 6. DenseNet architecture.

InceptionV3: Inception V3 (Fig. 7) model containing five convolution layers, one average pooling layer, two max-pooling layers, one fully related layer and 11 inception units which designed an image-wise classification (Vijayan, 2020).

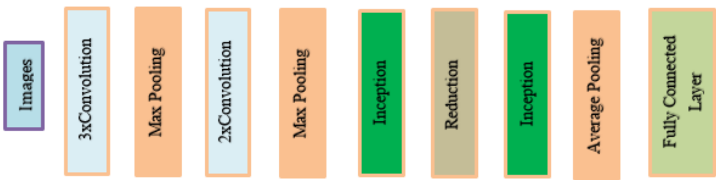


Figure 7. DenseNet architecture.

VGG-16: The name VGG-16 comes from the fact that it has 16 layers. It contains different layers, including convolutional layers, activation layers, max-pooling layers, and fully connected layers.

There are 13 convolutional layers, 5 max-pooling layers, and 3 dense layers. These amount to 21 layers in total, but there are only 16 weight layers (Fig. 8) (Hariri, 2022).

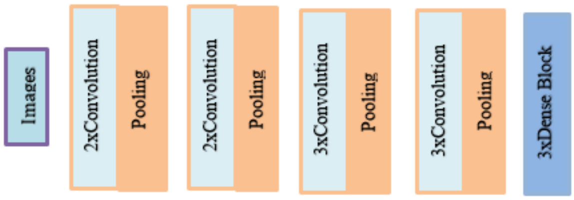


Figure 8. VGG-16 architecture.

MobileNet: MobileNet has two types of blocks and both blocks have three layers (Fig. 9). The first of these is 1x1 convolutions with ReLU. The second layer contains convolution and the third layer contains a non-linear 1x1 convolution. The first layer is a one-step residual block. Except for the last fully connected layers, most backbone networks for sensing are networks for classification tasks. The backbone network produces feature maps for each input image using the images it receives as input (Almghraby, 2021).

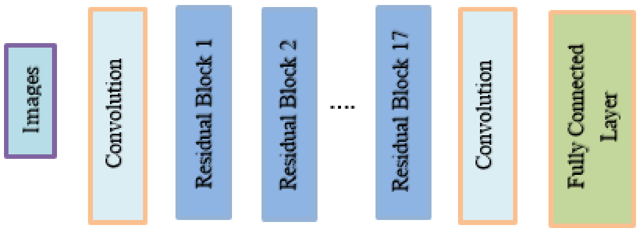


Figure 9. MobileNet architecture.

NasNetMobile: NasNet (Fig. 10) is a CNN architecture consisting of simple blocks containing basic cells, improved by reinforcement training, and separable convolution and concatenation. NasNet-based architectures are formed according to the logic of repeating these simple blocks according to the capacity of their networks. NasNetMobile, which has 769 layers and an input size of 224x224, consists of 12 cells with a capacity of 5.4 million and a multiplying capacity of 564 million. It consists of multiple replicated access volumes, each with the same architecture and different weights. If a feature is mapped to create a simply scalable architecture for images of any scale, the convolutional unit has the role of returning the feature map of the same size. Additionally, it has two important roles in rotating the feature map, which is twice the height and width of the feature map (Daşgın et al., 2023).

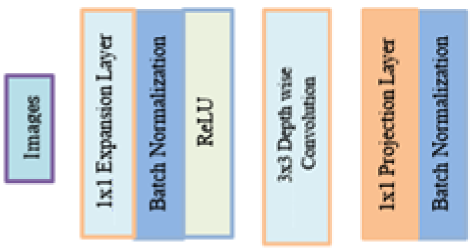


Figure 10. NasNet architecture.

Xception: The Xception model is based on depth separable convolution, which divides ordinary convolution into two procedures: point convolution and spatial convolution. Spatial convolution is performed independently on each input channel. Then pointwise convolution uses a 1×1 kernel for point-to-point convolution. It reduces both the number of parameters and calculations. It contains 14 residual blocks, a total of 3 joint convolution layers and 33 depth-separable convolutions. All three co-convolution layers are in the pre-processing module. However, Xception uses global average pooling with a full-connection layer in its decision module (Chen, 2021).

Xception, one of the networks used, contains deeply separable convolution layers. Figure 11 shows the network structure of the Xception network. As seen in Figure 11, at the beginning of the network, two convolution layers with a filter width of 3x3 and the outputs of the convolution layer with a filter width of 1x1 were combined with the outputs. Then the state of the data before entering the convolution block was combined with three convolution layers with a filter width of 3x3 and this process was repeated 8 times.

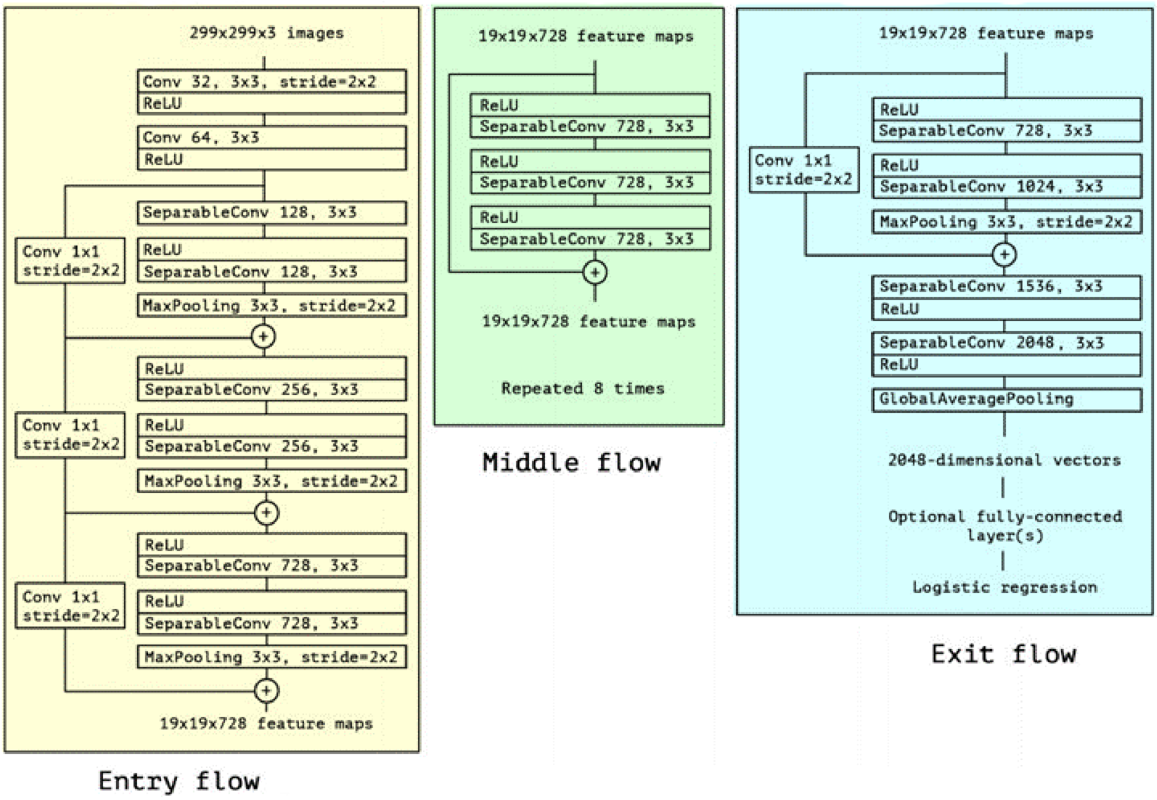


Figure 11. Network structure of the Xception network [URL 3].

The F1 performance measurement score was used to evaluate the performance of the methods used. The F1 score provides a score that balances both precision and recall concerns in a single measurement using the following equation (Snyder and Husari, 2021):

$$F_1 = \frac{2x(precisionxrecall)}{recisionxrecall} \quad (5)$$

3. RESULTS

A total of 2647 images were used in the data set, 2397 of which were used for training and 250 for testing. In 223 of them masks in the images are coloured in computer to train the model with coloured masks. These images are trained in different convolutional neural networks and their accuracy is compared. Traditional convolutional networks and pretrained networks were used. Traditional convolutional networks include two networks, four-layer and eight-layer, pre-trained networks ResNet-50, DenseNet-121, EfficientNet-B3, VGG-16, MobileNet, NasNetMobile, Xception networks. Random rotation, flip and translation were used for data augmentation. 500 x 750 x 1 input resolution is used in Traditional convolutional networks, 400 x 400 x 3 input resolution is used in pre-trained networks (224x224x3 input resolution is used in NasNetMobile network).

Most pre-trained networks performed better than sequential networks. However, some pre-trained networks performed very poorly. In Table 4, the accuracy and loss in the epoch where each network gives the highest accuracy and the number of this epoch are given after training the top of the network.

The highest accuracies were obtained in MobileNet-v2 and Xception networks. An additional fully connected layer with 512 neurons is added to networks before the classification layer. At this stage, different learning rates were tried and the best results in the training phase were obtained with a learning rate of 0.001. The top of the networks are trained 10 epochs with 0,001 learning rate and the networks are fine-tuned 20 epochs with 0,0001 learning rate. MobileNet-v2 network overfitted quickly and fine-tuned Xception network gave better results. The results of fine-tuned networks in the epoch where each network gives the lowest loss are given in Table 5.

Table 4. Results obtained according to network structures.

Model	Val. loss	Val. accuracy	Val. F1 score	Val. precision	Trai. loss	Trai. accuracy	Trai. F1 score	Trai. precision	Epochs
MobileNet-v2	0.2155	0.944	0.9426	0.9536	0.2509	0.9145	0.9165	0.9375	18
Xception	0.1784	0.936	0.9345	0.9435	0.1970	0.9344	0.934	0.9426	10
DenseNet-121	0.2014	0.932	0.9317	0.939	0.2352	0.9334	0.9309	0.9514	25
Inception-v3	0.243	0.932	0.9312	0.9339	0.2114	0.9383	0.9364	0.9578	16
VGG-16	0.257	0.904	0.902	0.9253	0.3017	0.8962	0.8973	0.9192	19
NasNetMobile	0.3758	0.86	0.8609	0.8866	0.4143	0.8516	0.8496	0.8972	20
8 layer CNN	0.541	0.808	0.8102	0.8689	0.5938	0.7748	0.773	0.84	25
4 layer CNN	0.8169	0.712	0.7092	0.8314	0.8435	0.6861	0.6825	0.8034	20
ResNet-50	1.3858	0.396	0.3816	0.4917	1.6592	0.3665	0.3531	0.4226	15

Table 5. Results of fine-tuned models

Model	Val. loss	Val. accuracy	Val. F1 score	Val. precision	Trai. loss	Trai. accuracy	Trai. F1 score	Trai. precision
Xception	0.0586	0.98	0.9794	0.98	0.1428	0.9512	0.9503	0.9597
MobileNet-v2	0.43	0.85	0.8192	0.8383	0.3742	0.8598	0.8531	0.8705

Fig. 12 shows training and validation metrics (accuracy and precision) of training (epochs 0-10) and fine tuning (epochs 10-30). Losses are shown in Fig. 13.

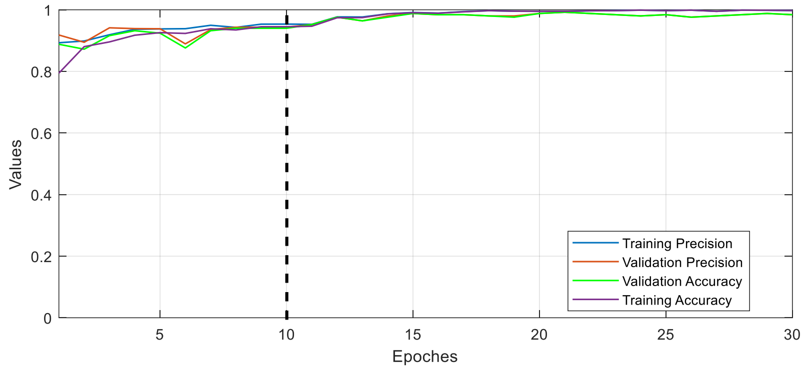


Figure 12. Accuracy and precision plots.

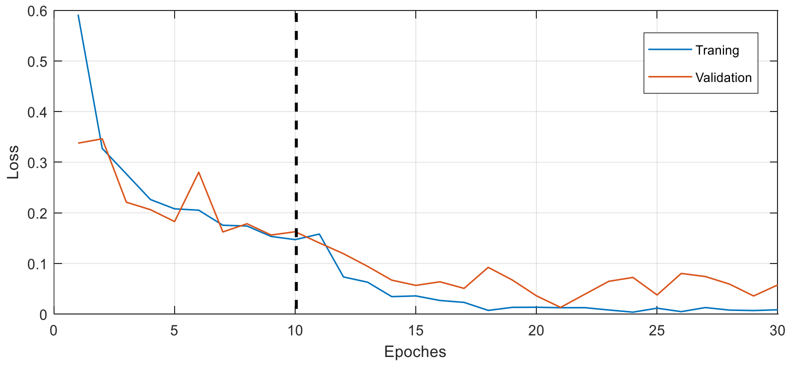


Figure 13. Loss plots.

4. CONCLUSION

The importance of determining mask quality is increasing after the COVID-19 pandemic. In this study, a software that will determine whether the masks are reliable or not has been implemented using Python. Five different types of masks were classified using different network structures and different convolutional neural network models. For this purpose, a database consisting of 2424 mask photographs has been prepared. Random flip and rotation were used for data augmentation.

Four Layer CNN, Eight Layer CNN, ResNet-50, DenseNet-121, EfficientNet-B3, VGG-16, MobileNet, NasNetMobile and Xception CNN models were used to classify masks and the results were compared. The highest accuracy is achieved by adding a fully connected layer to the Xception network. 223 images are shot and masks in the images are coloured to train the new model with coloured masks. The type of mask was correctly determined with 99.2% accuracy. In the data set prepared with the intended method, mask photographs were taken from close range. The system may be less accurate for photos taken from long distances.

With this developed study, the user can be informed whether the mask is reliable or not. In future studies, the model can be turned into a mobile application where users can test masks before purchasing them.

Conflict of interest
The authors declare that they have no conflict of interest.

REFERENCES

Almghraby, M. and Elnady, AO. (2021). Face Mask Detection in Real-Time using MobileNetv2. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-10 Issue-6, pp: 104-108.

Asif, S., Wenhui, Y., Tao, Y., Jinhai, S. and Amjad, K. (2021). Real Time Face Mask Detection System using Transfer Learning with Machine Learning Method in the Era of Covid-19 Pandemic. 4th International Conference on Artificial Intelligence and Big Data (ICAIBD), 70-75. Doi: 10.1109/ICAIBD51990.2021.9459008.

Bozkurt, F. (2022). A Comparative Study on Classifying Human Activities Using Classical Machine and Deep Learning Methods. Arabian Journal for Science and Engineering 47:1507–1521. Doi: doi.org/10.1007/s13369-021-06008-5

Cahvda, A., Dsouza, J., Badgujar S. and Damani A. (2021). Multi-Stage CNN Architecture for Face Mask Detection. 6th International Conference for Convergence in Technology (I2CT) Pune, India. Doi: 10.1109/I2CT51068.2021.9418207.

Chen, B., Ju, X., Xiao, B. et al. (2021). Locally GAN-generated face detection based on an improved Xception. Information Sciences Volume 572, September 2021, Pages 16-28.

Ciuffreda, S., Picotti, C., Pescio, P. (2021). Medical face masks on the market: Review of materials, characteristics and performed tests. Medical Device Testing. Eurofins Biolab.

Daşgın, A., Adem, K. & Kılıçarslan, S. (2023). Detection of Face Mask with Convolutional Neural Network Models to Reduce Covid19 Spread. Journal of the Institute of Science and Technology, 13(3): 1511-1527.

Du, X., Cai, Y., Wang, S. and Zhang, L. (2016). Overview of deep learning. 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). 159-164, Doi: 10.1109/YAC.2016.7804882.

Fırat, H., Hanbay, D. (2023). Comparison of 3D CNN based deep learning architectures using hyperspectral images, Journal of the Faculty of Engineering and Architecture of Gazi University, 38:1, 521-534.

Goyal, H., Sidana, K., Singh, C., et al. (2022). A real time face mask detection system using convolutional neural network. Multimed Tools Applications. 81, 14999–15015 Doi: 10.1007/s11042-022-12166-x.

Hariri, W. (2022). Efficient masked face recognition method during the COVID-19 pandemic. SIViP 16, 605–612. https://doi.org/10.1007/s11760-021-02050-w.

Hasan, N., Bao, Y., Shawon, A. et al. (2021). DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image. SN COMPUT. SCI. 2, 389. https://doi.org/10.1007/s42979-021-00782-7.

Kansal, I., Popli, R. and Singla C. (2021). Comparative Analysis of various Machine and Deep Learning Models for Face Mask Detection using Digital Images. 9th International Conference on Reliability, Infocom Technologies and Optimization. Doi: 10.1109/ICRITO51393.2021.9596407.

Kayalı, D., Dimililer, K., Sekeroğlu, B. (2021). Face Mask Detection and Classification for COVID-19 using Deep Learning. International Conference on INnovations in Intelligent SysTems and Applications (INISTA). 10.1109/INISTA52262.2021.9548642.

Kaur, G., Sinha, R., Tiwari, PK., Yadav, SK., Pandey, P., Raj, R., Vashisth, A., Rakhra, M. (2022). Face mask recognition system using CNN model. Neuroscience Informatics, 2, 3, 100035. Doi: /10.1016/j.neuri.2021.100035.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, DJ. (2021). 1D convolutional neural networks and applications: A survey, Mechanical Systems and Signal Processing. 151, 107398. Doi: 10.1016/j.ymssp.2020.107398.

Köklü, M., Çınar, I., Taşpınar, YS. (2022). CNN-based bi-directional and directional long-short term memory network for determination of face mask. Biomedical Signal Processing and Control. 71. Doi: 10.1016/j.bspc.2021.103216.

Naufal, MF. et al. (2021). Comparative Analysis of Image Classification Algorithms for Face Mask Detection. Journal of Information Systems Engineering and Business Intelligence. Doi: 10.20473/jisebi.7.1.56-66.

O'Kelly, E., Arora, A., Pirog, S., Ward, J., Clarkson, PJ. (2021). Comparing the fit of N95, KN95, surgical, and cloth face masks and assessing the accuracy of fit checking. Plos One. 16(1): e0245688. Doi:10.1371/journal.pone.0245688.

O'Shea, K., Nash, R. (2015). An Introduction to Convolutional Neural Networks. arXiv preprint arXiv: 1511.08458.

Rajath, AN., Shruthi, BM., Ambareen, K. and Lakshmi, CMS. (2021). An Adaptive Approach to Detect Face Mask in Real Time using Convolutional Neural Network (CNN) Model, 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, 826-831. Doi: 10.1109/ICEECCOT52851.2021.9708046.

Rokhana, R., Herulambang, W., Indraswari, R. (2021). Multi-Class Image Classification Based on MobileNetV2 for Detecting the Proper Use of Face Mask. International Electronics Symposium. Doi: 10.1109/IES53407.2021.9594022.

Sighencea, BI., Stanciu, RI., C˘aleanu, CD. (2021). A Review of Deep Learning-Based Methods for Pedestrian Trajectory Prediction. Sensors, 21, 7543. Doi: 10.3390/s21227543.

Snyder, S. and Husari, G. (2021). Thor: A Deep Learning Approach for Face Mask Detection to Prevent the COVID-19 Pandemic. Southeast Conference. Doi: 0.1109/SoutheastCon45413.2021.9401874.

Sethi, S., Kathuria, M., Kaushik, T. (2021). Face mask detection using deep learning: An approach to reduce risk of coronavirus spread. Journal of Biomedical Informatics. 120, Doi: 103848. doi.org/10.1016/j.jbi.2021.103848.

Sharma Y., Mishra M., Furqan A. (2022). Face Mask Detection Using IoT and Deep Learning for Saftey of Covid-19. International Journal of Engineering, Science, Technology and Innovation (IJESTI), 2, 3.

Song, Z, Nguyen, K, Nguyen, T, Cho, C and Gao, J. (2022). Spartan Face Mask Detection and Facial Recognition System. Healthcare. 10, 87. Doi: 0.3390/healthcare10010087.

Su, X, Gao, M, Ren, J, Li, Y, Dong, M, Liu, X. (2022). Face mask detection and classification via deep transfer learning. Multimed Tools Appl. 81(3): 4475–4494. Doi: 10.1007/s11042-021-11772-5.

Teboulbi, S., Messaoud, S., Hajjaji, MA. and Mtibaa, A. (2021). Real-Time Implementation of AI-Based Face Mask Detection and Social Distancing Measuring System for COVID-19 Prevention. Scientific Programming. Article ID 8340779, 21. 10.1155/2021/8340779.

Tomás, J., Rego, A., Viciano-Tudela, S. and Lloret, J. (2021). Incorrect Facemask-Wearing Detection Using Convolutional Neural Networks with Transfer Learning. Healthcare. 9, 1050. Doi: 10.3390/healthcare9081050.

URL 1, https://www.cordeks.com/Page.php?a=pp-meltblown, (Accessed: 21.05.2022).

URL 2, https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-overview/ (Accession date: 21.05.2022).

URL 3, https://maelfabien.github.io/deeplearning/xception, (Accessed: 21.06.2022).

Vijayan, T., Sangeetha, M., Karthik, B. (2020). Efficient Analysis of Diabetic Retinopathy on Retinal Fundus Images using Deep Learning Techniques with Inception V3 Architecture. Journal of Green Engineering (JGE) Vol 10, 10.

Wakarekar, MM., Gurav, U. (2022). Image Processing and Deep Neural Networks for Face Mask Detection. In: Rajagopal, S., Faruki, P., Popat, K. (eds) Advancements in Smart Computing and Information Security. ASCIS 2022. Communications in Computer and Information Science, vol 1760. Springer, Cham. Doi: 10.1007/978-3-031-23095-0_14