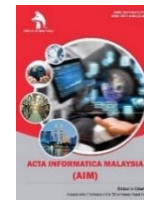




ZIBELINE INTERNATIONAL™  
P U B L I S H I N G  
ISSN: 2521-0874 (Print)  
ISSN: 2521-0505 (Online)  
CODEN: AIMCCO



## REVIEW ARTICLE

## A DEEP LEARNING MODEL FOR FACE RECOGNITION IN PRESENCE OF MASK

Kalembo Vikalwe Shakrani\*, Ngonidzashe Mathew Kanyangarara, Prince Tinashe Parowa, Vibhor Gupta, Rajendra Kumar

Sharda University, Greater Noida, India

\*Corresponding Author E-mail: 2020801862.kalembovikalwe@pg.sharda.ac.in

This is an open access journal distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

## ARTICLE DETAILS

## Article History:

Received 21 February 2022  
Accepted 25 March 2022  
Available online 04 April 2022

## ABSTRACT

Image classifications and object detection are common study topics in the rapidly expanding technological advancements to identify and detect real-time problems in major federal fields like public places, airports and army bases using webcams and surveillance cameras opensource platforms. The goal of this study is to suggest Open Source Computer Vision (OpenCV) and Convolutional Neural Network (CNN) techniques for identifying a person in presence of face mask from image datasets and real-time (live streaming video). For experimental purpose a parent directory consisting of three main directories (*i.e.*, training, testing and validation sets) and two sub directories inside those containing Mask (M) and No Mask (N), respectively are used. Mask subdirectories have images of people wearing masks and the vice versa is for Non Mask. Total 1006 images are used including 503 Mask and 503 No-Mask. The data augmentation pre-processing method is used to increase the dataset size to improve the accuracy of the suggested model. The proposed system uses a camera inbuilt on drone to capture real-time image for recognition using Conventional Neural Network (CNN). The proposed model is constructed, compiled and trained using Tensor flow and Keras. The final training accuracy recorded is 0.93, while the validation accuracy recorded is 0.94, the training loss is 0.17, the validation loss here observed is 0.1672, and the test loss is 0.15. The classification accuracy of the proposed system observed is 0.95.

## KEYWORDS

Conventional Neural Network (CNN), Open Cv, Pre Processing, Keras and Tensorflow.

## 1. INTRODUCTION

The coronavirus pandemic has largely impacted our daily lives, interfering with international business and movement. Putting on a protective face mask has become the norm for most people. In the not-too-distant future, most public service providers will require their clients to dress appropriately to use their services (Irfan et al., 2021). As a result, the detection of face masks has emerged as a critical assignment to assist society worldwide. Using some elementary artificial intelligence packages such as Keras, TensorFlow, Scikit-Learn, and OpenCV, this paper provides a simple approach to achieving this goal. The suggested procedure correctly identifies the face in the image and then determines whether or not a mask covers it in the image.

Surveillance cameras, webcams, and open-source platforms are used to identify and detect real-time issues in major public places and army locations. Image classification and object detection are common research topics in the rapidly growing field of advanced technology. With this research task, we suggest Open Source computer Vision (OpenCV) and Convolutional Neural Network (CNN) to identify coronavirus face masks from image datasets and live streaming video to determine COVID-19 face mask detection (Islam et al., 2020). Various techniques have been used for data augmentation pre-processing to increase the size of our dataset and improve the model's performance (Kumar et al., 2021). In addition, the MobileNet CNN architecture has been used to reduce the size of the dataset and the complexity of the models that have been proposed.

For this study, a drone's camera is used to track an object's journey from its origin to its final destination using a single-shot detector framework

and MobileNet, a quick and effective advanced learning-based technique for detecting objects in images. Using image processing and artificial intelligence, the coronavirus face mask detection model can identify people wearing or not wearing masks.

Face detection is a common research topic today, with various procedures being used and proposed an Arduino-based autonomous face detection system to determine real-time computer vision to ensure the safety of the deep learning security system. Using a Pan-Tilt mechanism with Arduino UNO is achieved using an ATmega328p Microcontroller clarifying procedures with Haar classifier Cascade and Open Source Computer Vision algorithm for image identification in real-time streaming data (Li et al., 2018). To create a safety system in required fields, it was proposed to use a high-speed image pre-processing technique. With the help of a webcam, this study uses a convolutional neural network and feature extraction to categorize and identify whether an individual is putting on a mask or not. Three levels of work are carried out to complete this study: the first step involves image pre-processing to bring out refined image, the next step requires image cropping (region of interest), and the final step involves image analysis using image categorization algorithms.

## 2. LITERATURE REVIEW

Image processing and object detection are increasingly being manipulated and detected using artificial intelligence and advanced learning methods in today's world and fast technological advancements. Several training and no-training based methods are developed for finger, face, vein recognition (Kumar et al., 2021; Kumar et al., 2019; Jugal and Rajendra, 2010; Vaibhav et al., 2016). In recent years, object detection for quadcopter drones was

## Quick Response Code



## Access this article online

Website:  
[www.actainformaticamalaysia.com](http://www.actainformaticamalaysia.com)

DOI:  
10.26480/aim.02.2022.38.41

proposed using the CNN deep learning model to detect the image faster. In the face recognition technique, a face is identified from an image that contains different features. The study of image recognition requires face tracking, pose estimation, and expression recognition techniques. The challenge is recognizing the person in a solitary image given only a face image. Face recognition is a challenging task because images vary in colour, shape, size, and other characteristics over time and are not immutable in any way. It becomes a time-consuming task when an opaque image is obstructed by something else that is not in front of the camera, and so on. This study presents occlusive face detection faces two significant challenges: the lack of readily available large datasets with both unmasked and masked images and the omission of facial expressions in the covered sector. Several misplaced manifestations can be recovered (Oumina et al., 2020).

The dominance of facial cues can be significantly reduced using dictionaries. The locally linear embedding (LLE) instructed an enormously massive collection of faces with masks and manufactured mundane images. The directories instructed on an enormously massive collection of faces with masks. Following the research published, CNNs in computer visualization have a severe setback on the input image size that must be

observed. To overcome the inhibition, it is common to reconfigure the images before incorporating them into the network. It is necessary to correctly identify the face in the image then recognize whether or not it is putting on a mask to complete the task successfully (Islam et al., 2020). The suggested technique should also recognize a face and a moving mask in the background to carry out surveillance tasks.

### 3. RESEARCH METHODOLOGY

The use of image processing and object detection is critical for most business enterprises to identify security problems and image categorization, particularly in health image processing. The essential motivation for this study is to suggest a mask identification system that uses the real-time computer vision (OpenCV) in conjunction with the CNN (an advanced learning algorithm) to identify the live streaming image in real-time (real-time) (Budiharto et al., 2018). As illustrated in figure 1, the classification findings are displayed regardless of whether the individual is putting on a mask or not. The general concept of the proposed model is that in the first phase, a random image dataset is collected that is freely available online and prepare the dataset, which is then uploaded to the server.

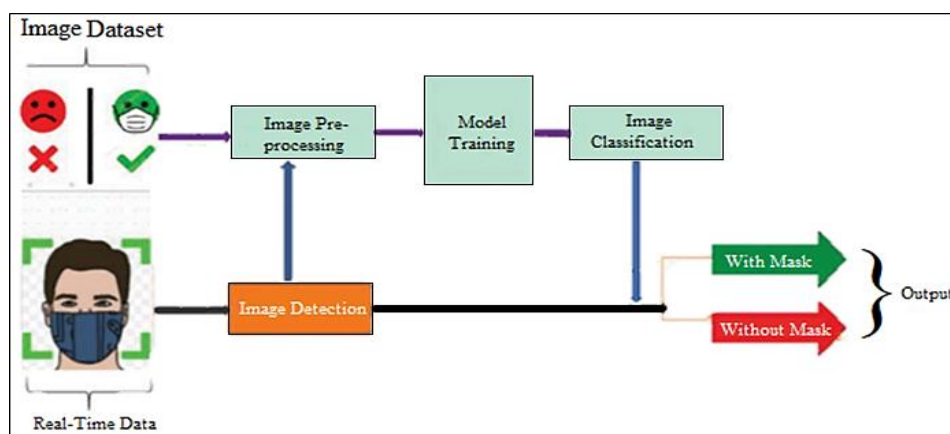


Figure 1: Image Classification Steps

Following the dataset's creation, the subsequent step in this area is introducing image pre-processing to retrieve elements and advance the model's accuracy (Thirupathi et al., 2021). After pre-processing, the model must be trained and predict the final output, which takes time as per

hardware configuration. The second step involves identifying real-time computer vision data using an OpenCV-built library and, lastly, displaying the results to determine if or not an individual is putting on a mask, as depicted in figure 2.

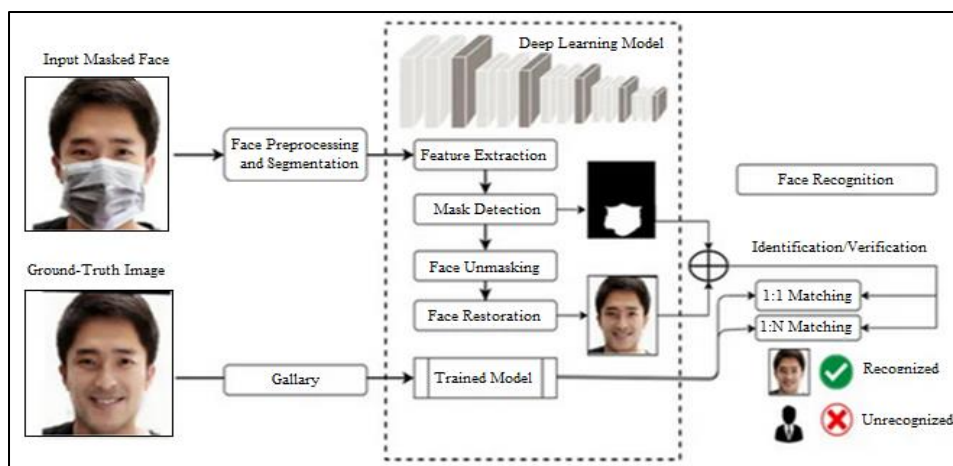


Figure 2: Model Training and Image Classification

#### 3.1 Data Augmentation and Image Processing

Pre-processing data helps in cleaning data, extracting essential features from a dataset, and improving model performance in classification. The image pre-processing is critical for improving image data and enhancing image aspects for further analysis. Image pre-processing helps to move data, brighten images, and create new data with lesser image datasets with an augmentation algorithm (Islam et al., 2020). The gathered classified data may vary in size, contrast, position, and orientation. Therefore image pre-processing is the deterministic, algorithm-based action taken to ensure the images are correctly edited and ready for training and inference. Augmentation techniques are used to pre-process the image dataset. Using ImageGenerator and OpenCV built-in functions, It can

randomly crop, colour filter, position, flip, and add noise to the training data to eliminate the overfitting of the model (Oumina et al., 2020). This study prepared the dataset using other pre-processing methods and augmentation to attain a verified dataset of various types. As shown in the figure 2, system passes preprocessed images to any advanced learning algorithms to retrieve features. Several data augmentation methods are applied to make sure we increase the size and quality of the dataset. This helped in reducing over fitting problems and enhances the model's overview ability during training. The settings deployed in image augmentation are shown in Table 1.

So firstly, the rescaling of the data is done by the factor of  $1/255$ . That helps in the normalization. After this the zoom range of 20%, is applied to

allow to take some random crops from the image. Rotation range is set to 40 degrees and horizontal flip is set to true.

### 3.2 Proposed CNN

CNN is an early application of image processing tasks such as text classification and object detection. The Convolutional Neural Network algorithm is superior to other advanced learning algorithms when extracting features from a dataset for image classification or object detection (Mengistie and Kumar, 2021). The ReLu activation function and convolutional filters remove elements from the Convolutional Layer. Convolutional neural networks (Conv2D) have been used extensively for image classification, but only one dimension has been used for text classification based on this principle (Islam et al., 2020). This model is created for human identification in presence/absence of face mask, and it is modified to fit the issues in the domain. The significant elements of this Convolutional Neural Network model are depicted below.

### 3.3 Input Image

The accumulated input image is set for pre-processing to make the model's execution effective. To maximize the datasets, data augmentation techniques are used. Physical data collection is complex due to the COVID19 global pandemic (Oumina et al., 2020). The augmented data should be sent to the preceding convolution to retrieve essential elements.

### 3.4 Convolution steps

Convolutions extract local aspects from large input data sets and multiply the resulting  $N \times N$  matrices. Conv2D use filters, kernel size, input shapes and activation functions for image classification (Mengistie and Kumar, 2021). In this study domain, the values of the variables are already stated in table 1.

### 3.5 Max Pooling Layer

In advanced learning algorithms, max-pooling reduces dimensionality and extracts maximum features. The pooling layer reduces the number of variables and regularizes overfitting by finding the average of the provided elements (Oumina et al., 2020).

### 3.6 Fully Connected Layer

It plots pooling layers, flattened and fed into the subsequent layer. The fully connected layer is essential for classification in CNN. The following classification outcome is performed with the aid of the activation function (sigmoid, Dense). The summary of layers of the proposed network architecture is shown below.

### 3.7 Model: "sequential"

Layer (type)	Output Shape	Param #
Conv2d (Conv2D)	(None, 150, 150, 32)	896
max_pooling2d	(None, 75, 75, 32)	0
Dropout (Dropout)	(None, 75, 75, 32)	0
conv2d_1 (Conv2D)	(None, 75, 75, 64)	18496
max_pooling2d_1	(None, 37, 37, 64)	0
dropout_1 (Dropout)	(None, 37, 37, 64)	0
Flatten (Flatten)	(None, 87616)	0
Dense (Dense)	(None, 256)	22429952
dropout_2 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257

Total params: 22,449,601

Trainable params: 22,449,601

Non-trainable params: 0

## 4. RESULT

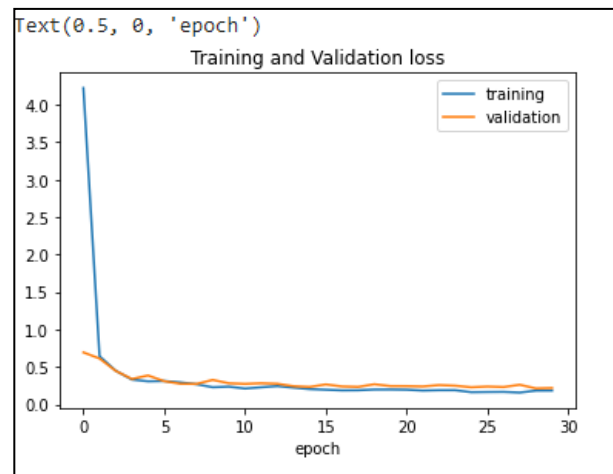
In the proposed model Adam Optimizer with a learning rate of 0.001 is compiled. The model is trained with all the training images with 30 epochs. During training of the model, some values per epoch were lost namely accuracy, validation loss and validation accuracy. So the loss and accuracy is an indication of progress of the training. It made a guess that the classification of the training data, and then measuring it against the new label.

While the accuracy is the percentage of correct guesses, and the validation accuracy is the measurement with the data that it has not been used in

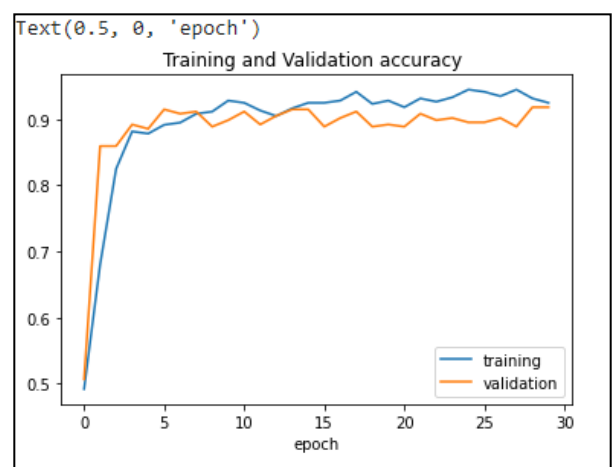
**Table 1:** Image Augmentation Settings

Method	Setting
Rescale	1/255
Zoom Range	0.2
Rotation_range	40
Horizontal Flip	True

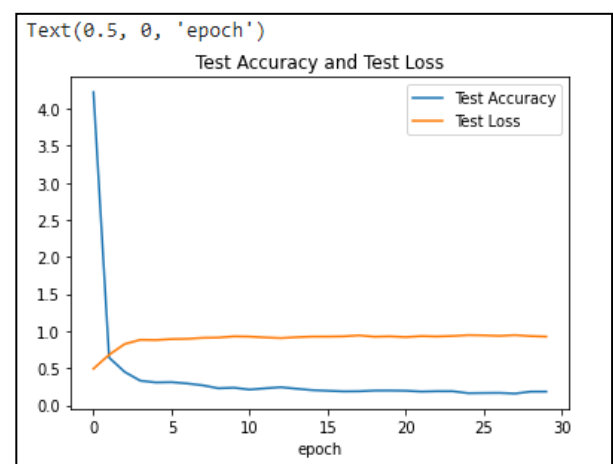
training. After training, two graphs are plot showing the accuracy of training, one for a loss, while the other one for the accuracy. The final training accuracy recorded is 0.93, while the validation accuracy observed is 0.94, the training loss is 0.17, the validation loss is 0.1672. The classification accuracy observed is 0.95 and the test loss is 0.15. The training results are presented by figures 3-5.



**Figure 3:** Training and Validation Loss



**Figure 4:** Training and Validation Loss



**Figure 5:** Test Accuracy and Test Loss

## 5. CONCLUSION

In conclusion, due to the global pandemic caused by COVID-19, many countries are employing a variety of mechanisms and innovations to combat the virus. This group of mechanisms includes protection mechanisms such as a face mask. While working on the coronavirus face mask detection and classification algorithm OpenCV, the central focus of this study is to determine (classify) if or not an individual is putting on a face mask with human identification. The Convolutional Neural Network deep learning categorization algorithm OpenCV is used to accomplish this research.

## REFERENCES

- Aiman, U., Vishwakarma, V. P. 2017. Face recognition using modified deep learning neural network, *8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-5
- Budiharto, W., Gunawan, A. A., Suroso, J. S., Chowanda, A., Patrik, A., Utama, G. 2018. Fast object detection for quadcopter drone using deep learning. In *Proceedings of 3rd International Conference on Computer and Communication Systems (ICCCS)*, pp. 192-195.
- Irfan, M., Akhtar, N., Ahmad, M., Shahzad, F., Elavarasan, R. M., Wu, H., Yang, C. 2021. Assessing public willingness to wear face masks during the COVID-19 pandemic: fresh insights from the theory of planned behavior. *International Journal of Environmental Research and Public Health*, 18(9), pp. 4577.
- Islam, M. S., Moon, E. H., Shaikat, M. A., Alam, M. J. 2020. A novel approach to detect face masks using CNN. In *Proceedings of 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, pp. 800-806.
- Jugal, K. G., Rajendra, K. 2010. An Efficient ANN Based Approach for Latent Fingerprint Matching, *International Journal of Computer Application, Foundation of Computer Science, USA, ISSN: 0975-8887*, 7(10), pp. 18-21.
- Kumar, R., Singh, R. C., Kant, S. 2021. Dorsal Hand Vein-Biometric Recognition using Convolution Neural Network, *Advances in Intelligent Systems and Computing*, pp. 1087-1107.
- Kumar, R., Singh, R.C., Sahoo, A.K. 2019. SIFT based Dorsal Vein Recognition System for Cashless Treatment through Medical Insurance, *International Journal of Innovative Technology and Exploring Engineering*, 8(10S), pp. 444-451.
- Li, Y., Zhang, M., Wang, W. 2018. Online Real-Time Analysis of Data Streams Based on an Incremental High-Order Deep Learning Model, in *IEEE Access*, 6, pp. 77615-77623.
- Mengistie, T. T., Kumar, D. 2021. Covid-19 Face Mask Detection Using Convolutional Neural Network and Image Processing, In *Proceedings of 2nd International Conference for Emerging Technology (INCET)*, pp. 1-7.
- Oumina, A., El Makhfi, N., Hamdi, M. 2020. Control the Covid-19 pandemic: Face mask detection using transfer learning. In *Proceedings of 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pp. 1-5.
- Thirupathi, J., Rao, N. K., Bala, M., Prathima, K., Naveen, D., Mallikarjun, G. 2021. Object Tracking Using OpenCV and Python. *Annals of the Romanian Society for Cell Biology*, 25(6), pp. 7815-7824.
- Vaibhav, J., Ajay, K. S., Rajendra, K. 2016. An Efficient Approach for Latent Fingerprint Recognition, *Proceedings of Second International Conference on Computational Intelligence & Communication Technology (CICT)*. <http://ieeexplore.ieee.org/document/7546577/>

