**International Journal of Innovative Research in Engineering**
*Volume 4, Issue 2 (March-April 2023), PP: 248-252.*
*www.theijire.com* **ISSN No: 2582-8746**

# Mask Detection & Alert with the Application of Open CV and Mobilenetv2

**Rathish Kumar M[1], Santhosh K V[2], Harish S[3], Hari Priya R[4]**
[1,2,3,4] *Computer Technology, Bannari Amman Institute of Technology, Tamilnadu, India.*

**How to cite this paper:**
Rathish Kumar M[1], Santhosh K V[2], Harish S[3], Hari Priya R[4], "Mask Detection & Alert with the Application of Open CV andMobilenetv2", IJIRE-V4I02-248-252.



***Abstract:*** *Face mask detection with an alert system is a technology designed to automatically detect whether individuals are wearing face masks in public spaces.This system uses computer vision and deep learning algorithms to analyze video feeds from cameras and identify whether individuals in the footage are wearing masks or not. When the system detects individuals not wearing masks, it triggers an alert mechanism, such as a warning message via email to the admin, to prompt the individuals to wear a mask. Face mask detection with an alert system has become increasingly relevant in the context of the COVID-19 pandemic and otherspread of viruses, as it can help enforce mask-wearing policies and reduce the spread of the virus in public spaces. The use of face masks has become essential in any outbreak like the COVID-19 pandemic in the future and also we can use face mask detection in chemical industries for the safety of workers by wearing masks.*
***Key Word:*** *Face mask detection, Convolutional Neural Network (CNN), Mobile Net V2 model.*

## I.INTRODUCTION

The fundamental objective of face mask detection is to categorize the masked and unmasked images or videos of people to give the percentage of coverage level of people wearing face masks or not wearing masks.

**Introduction about face mask detection:** The COVID-19 pandemic has changed the way we live, work, and interact with others. One of the most important measures to prevent the spread of the virus is wearing a face mask. However, ensuring that individuals comply with mask-wearing regulations in public spaces can be challenging. Face mask detection using computer visionand machine learning algorithms is a promising solution to this challenge.

Convolutional Neural Networks (CNNs) is a type of deep learning algorithm that is commonly used for image classification tasks. They are well-suited for tasks such as face mask detection due to their ability to learn features from raw pixel data.To develop a CNN for face mask detection, the first step is to obtain a large dataset of images containing faces with and withoutmasks. This dataset is used to train the CNN to recognize the presence or absence of a face mask in an image.

The CNN typically consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers.The convolutional layers perform feature extraction by applying filters to the input image to identify important patterns and features.The pooling layers downsample the output of the convolutional layers to reduce the dimensionality of the feature maps. Finally, thefully connected layers perform classification by applying weights to the features and making a prediction on whether the input imagecontains a face mask or not.

One popular CNN architecture for face mask detection is the MobileNetV2 model, which is optimized for mobile devices and has a small footprint. This model is trained on a large dataset of images containing faces with and without masks and then is deployed in real-time to detect faces and classify them by wearing a face mask or not wearing a face mask. Overall, CNNs are an effective tool for face mask detection due to their ability to learn complex features from images and make accurate predictionsin real time.

**Role of convolutional neural network:** Convolutional Neural Networks (CNNs) play a critical role in image processing and computer vision applications, including face mask detection. The key role of CNNs is to automate the feature extraction process from raw image data, which is often too complex and time-consuming for humans to perform manually.

In traditional machine learning algorithms, the feature extraction process requires human domain expertise to identify relevant features from raw data. However, CNNs automatically learn features from the raw image data through convolutional layers, which perform a series of convolution operations on the input image. These convolution operations identify patterns in the image data, such as edges, corners, and other relevant features, and then pass them on to the next layer for further processing. Pooling layers are another important component of CNNs. They reduce the dimensionality of the feature maps generated by the convolutional layers by down sampling them. This reduces the number of parameters that the network needs to learn and reduces the likelihood of over fitting. The final layer of a CNN is typically a fully connected layer that performs classification. This layer takes the features extracted by the convolutional and pooling layers and applies weights to them to makea prediction about the image class.

In the context of face mask detection, CNNs can be used to detect faces in an image and then classify them as either wearing or not wearing a face mask. This is achieved by training the CNN on a dataset of images containing faces with and without masks. The CNN learns to distinguish between these two classes by extracting relevant features from the images.

Overall, the role of CNNs in image processing and computer vision is to automate the feature extraction processand make accurate predictions in real time. This makes them a powerful tool for tasks such as face mask detection, object recognition, and image classification.

## II.LITERATURE REVIEW

1. According to Mrs. Monali Nitin Chaudhari and co .In order to improve current Viola Jones model. They have added CNN to it. The improved model builds the precision of face identification to 90% where the existing algorithm is simply about 78.4%.Under Occlusions this algorithm also fails to detect faces [1].
2. In this paper they have proposed to build a face recognition model using Convolutional Neural Network. They have combined the convolution and sampling layers in order to simplify the cnn model. The benefit of this model is that the memory requirements are decreased, and therefore number of parameters to be trained is correspondingly decreased which also lessens the training time [2].
3. According to SukhadaChokkadi, Sannidhan MS, Sudeepa K B, AbhirBhandary they have conducted literature review on different face recognition algorithms and along with what is the output of those algorithms and also they have mentionedabout its limitations [3].
4. In this paper they are using CNN model , image super-resolution and SRCNet to detect face mask.The outcomes acquired were more precise yet quantities of datasets utilized were little and this model can't be tried on video transfer since they haven't added video in the dataset [4].
5. According to Vinitha.V and Velantina.V they have identified whether the individual is covering their face by mask or not with deep learning and computer vision which utilizes Python, OpenCV, TensorFlow and Keras .And also this paperclearly explains the above APIs and how it works [5].
6. Yan, K., Huang and his teammates are by utilizing Convolution Neural Network they have proposed a face recognition algorithm. They prepared the feature extraction and the classifier by using stochastic gradient decent method. If any over-fitting issue is discovered, they use dropout technique to address it. During testing and preparing measure, they have utilized Convolution Architecture for Feature Extraction structure. The acknowledgment rate acquired was 99.82% and 99.78% [6].
7. In this paper they have concocted the deep learning model along with ML methods is used for face mask identification. The proposed model comprised two stages feature extraction base on ResNet50 2) classification based on SVM, decision tree. Three different datasets were used as benchmarks to evaluate the proposed technique. The SVM based modelachieved most elevated accuracy of 99.64% [7].
8. In this paper they have proposed a model and dataset to track down the normal and masked face images in nature. They introduced a tremendous dataset Masked Faces (MAFA), which has 35, 806 mask photographs. The proposed model ward on a convolutional neural organization called LLE-CNNs, which includes three modules (a suggestion, introducing, and check). The works showed that LLE-CNNs using MAFA achieved the ordinary precision identical to 76.1% [8].
9. In this paper they have proposed an algorithm that is to detect and recognize faces. They are using PCA algorithm for detection faces adequately. Lastly they got a general accuracy for face identification is 98.8 % and for correct facial recognition the accuracy they got is 99.2% [9].
10. In this paper K. Sunil Manohar Reddy, made a review on different face recognition methods and also explained about each of its positives and negatives about the algorithm. He did the overview on some algorithm like neural networks, Elastic Bunch Graph Matching, Fisherface [10].

## III.CURRENT TECHNOLOGIES IN DEEP LEARNING

### Convolutional neural network

A convolutional neural network (cnn) is a type of artificial neural network that is commonly used for analyzingvisual imagery, such as images or videos. Cnns are specifically designed to recognize spatial patterns in data, making them well-suited for image classification, object detection, and other computer vision tasks.

The core component of a cnn is the convolutional layer, which applies a set of learnable filters to the input image to extract features at different spatial locations. These features are then fed into subsequent layers of the network, such as pooling layers and fully connected layers, to perform higher-level reasoning and classification.

Cnns are trained using a supervised learning process, where the network is shown a large number of labeled examples and learns to associate certain features with certain classes. During training, the network adjusts its internal parameters(weights and biases) to minimize the difference between its predicted output and the correct label for each example.

Cnns have achieved state-of-the-art performance on a wide range of computer vision tasks, including image classification, object detection, semantic segmentation, and more. They are widely used in industry and academia for a variety ofapplications, such as self-driving cars, facial recognition, and medical imaging.

### Computer vision

Computer vision is a field of artificial intelligence and computer science that focuses on enabling machines to interpret and understand visual data from the world around them, such as images or videos. The goal of computer vision is to develop algorithms and systems that can analyze, interpret, and extract meaningful information from visual data, similar to how humans perceive and understand the world.

Computer vision applications are numerous and diverse, ranging from basic image processing tasks like image enhancement and denoising to more complex tasks such as object detection, facial recognition, and autonomous driving. Commonapplications used in computer vision include:

- Object recognition and detection: identifying objects and their locations within an image or video stream, such as recognizing faces in photos or identifying vehicles in a traffic scene.
- Scene reconstruction and 3d modeling: creating 3d models of objects or environments from 2d images or video streams.
- Image and video analysis: analyzing visual data to extract information such as motion, color, texture, and shape.
- Augmented reality: overlaying digital information or graphics onto the real-world view captured by a camera.
- Medical image analysis: analyzing medical images to aid in the diagnosis and treatment of diseases.
- Robotics: enabling robots to perceive and navigate their environment using visual sensors.

Computer vision algorithms and techniques typically involve machine learning, image processing, pattern recognition, and other fields of computer science and mathematics. The advancements in computer vision have led to significant improvements in many areas of human life, such as healthcare, transportation, entertainment, and security.

**Keras**

Keras are often used together in building and training convolutional neural networks (CNNs).

Keras provides a high-level api that simplifies the process of defining and training CNNs. With keras, developers can quickly and easily define the architecture of a CNN by stacking layers, such as convolutional layers, pooling layers, and fully connected layers, on top of each other. Keras also provides a wide range of pre-built layers, loss functions, andoptimizers that can be easily used in CNNs.

Overall, keras is powerful tools for building and training CNNs, and their combination provides an efficientand flexible way to implement deep learning models for computer vision tasks.

**Tensor flow**

Tensor flow are often used together in building and training convolutional neural networks (CNNs). Under the hood, keras uses tensor flow as its backend to perform the computations required to train the CNN. Tensor flow provides a low- level API that allows developers to implement custom layers and operations, optimize computations for GPUs, and perform distributed training across multiple devices. By using keras on top of tensor flow, developers can take advantage of the ease of useand flexibility of keras, while also benefiting from the performance and scalability of tensor flow. Overall, tensor flow is powerful tools for building and training CNNs, and their combination provides an efficient and flexible way to implement deep learning models for computer vision tasks.

## IV.METHODOLOGY

The figure 1. shows the actual workflow of the project in which the image acquisition is the first phase where the inputs has been provided. Then pre processing the image by terminating the irrelevant items from the input. Extraction processis the prior process in which the feature has been extracted from the input. Then it enters into the CNN model where the face has been recognized with the help of trained dataset. Then recognizing the face whether the person covered his face with mask or not.If not the snap of the person has been delivered to the admin via email.

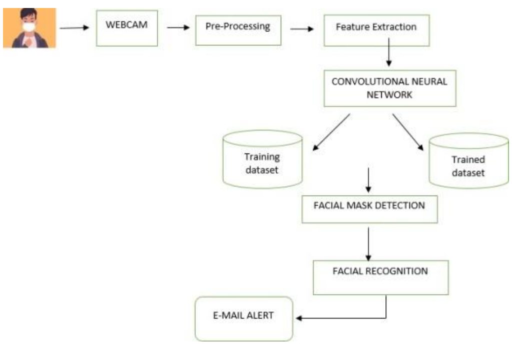


*Figure no 1 Workflow of the Project*

**Image Acquisition**

The input is provided by using webcam with a live video stream. The real time video is to be converted into frames per second so that it paves the way for the implementation of the face recognition to detect whether the subject wears the mask or not. Each frames from the video stream is considered for the comparison of the trained dataset to the frames.

**Dataset Collection**

Each dataset for the project has been collected from the kaggle which is well known for the collection of dataset. From kaggle the dataset with mask, without mask and improper mask. These datasets are properly preprocessed and implied to the process of training and testing process to undergo the detection of the mask.

**Preprocessing**

The dataset extracted from the kaggle is unstructured and raw so the dataset should undergo several processes to get more structured data. The first step is to define a standard image size to all the images from the dataset which advances the further process. Following that, we use the img to array function to transform the image into an array of data, and then append the data and labels to the list. Following that, numpy arrays are being converted. We are one-hot encoding the labels using label binarizer.The data is then divided into training and testing. 75% of the data are used for training, and the remaining 25% are used for testing. After that, we add data by utilising the image data generator function. We expand our dataset to provide a larger number of imagesfor the model's training.

**Face net**

The facial recognition algorithm known as Face Net was put forth by Google researchers in the paper Face Net: A Unified Embedding for Face Recognition and Clustering published in 2015. It produced cutting-edge results in numerous benchmark face recognition datasets, including Youtube Face Database and Labeled Faces in the Wild (LFW). It makes use of a neural network to train a mapping from face images to a high-dimensional space where the similarity of the related faces is determined by the Euclidean distance between two representations.

In order for Face Net to function, a face must first be identified and aligned in an input image. After that, the aligned face must be fed into a deep neural network, which then extracts a 128-dimensional embedding for that face. While comparing faces, the embedding is used to determine which faces are more similar to one another.

Face Net has been used in a range of applications, including security systems, social media tagging, and augmented reality, after achieving state-of-the-art performance on various facial recognition benchmarks. They suggested a method in which deep learning architectures like ZF-Net and Inception Network are used to extract a high-quality face mapping from the photos. Then, it trained this architecture using a technique known as triplet loss as a loss function.

**Activation Function**

Nodes with activation functions are found at the end of or between neural networks (layers). They control whether the neuron fires or not. As it affects the calibre of model learning, the choice of activation function at hidden layers as well as at the output layer is crucial. While Softmax is utilised for the output layer and determines probability distribution from a real number vector, ReLU is mostly employed for hidden layers. For problems involving multi-class categorization, the latter option is favored. Compared to the sigmoid and tanh functions, ReLU delivers superior performance and more widespread depthlearning. Five Conv2D levels with ReLu activation functions and a 3 x 3 filter, plus five Max-Pooling layers with a 2 x 2 filter make up the entire system. The fully connected layers are Flatten and Dense. The output layer's activation function is softmax.

**Mobile Net V2 Model**

Because MobileNetV2 from is employed in picture recognition and classification issues, we decided to use it when creating the mask detection model. 53 layers deep, it. Moreover, the MobileNetV2 model requires 30% less parameters than comparable models and is faster throughout the full latency spectrum. We first declared the base model to be MobileNetv2, and we are now passing the output of the base model to the head model. Following that, we used AveragePooling2d. After that, we converted the array of pixel values into a single vector using the Flatten function (3X3 into one single Vector). Dropout function to prevent model overfitting problems. and utilised dense function to appropriately return the output. We used softmax as the activation function so that our findings might be understood.

## V.RESULT AND DISCUSSION

The results of our model are shown in this section. The following metrics were taken from the input images foreach class (Mask / No Mask) in order to evaluate the results. Our suggested model achieves good results using a combination of Python bundles. In order to develop our model, I first imported all the necessary packages,

1. Matplotlib is what we're utilising for the visualisation.
2. Numpy for listing the values of the pixels
3. cv2 for image reading
4. Sklearn for model training and task preprocessing
5. As well as Tensor flow for creating the model,

The model's output appears promising. They are three people, as can be seen in figure 6. Person 1 is not wearingthe mask at all it predicted accurately and the person 2 is wearing mask correctly and our model predicted correctly and person 3is \spartially wearing the mask and her nose and mouth are not covered by mask and our \smodel predicted correctly that person is not wearing mask.

Table no 1. shows that both classes with and without masks have the same precision of (0.99), and both classes have the same recall (0.99). Moreover, the f1-score for each class is the same (0.99), with Class 1 receiving 433 data points of support and Class 2 receiving 386 data points. We may say that our model provides both the class and better, more accurate results (0 & 1).Our model's overall average is 99%. ( 0.99). It is simple to see how model accuracy varies between train and test datasets from the graphic below. According to our interpretation, the best accuracy for train data is roughly 0.97, which occurs after epoch 3. After epoch 4, it retains the same accuracy but may fluctuate somewhat. We can quickly see the model loss between the train and test datasets from the graphic above. We can take that as meaning that the data with the biggest loss for trains is between epochs (0-5).

**Table no 1:** Shows accuracy of the model

| | Precision | Recall | F1 score | Support |
|---|---|---|---|---|
| With mask | 0.99 | 0.99 | 0.99 | 433 |
| Without mask | 0.99 | 0.99 | 0.99 | 386 |
| Accuracy | | | 0.99 | 819 |
| Macro avg | 0.99 | 0.99 | 0.99 | 819 |
| Weighted avg | 0.99 | 0.99 | 0.99 | 819 |


## VI.CONCLUSION

In this suggested system, we have provided some initial results on mask identification and face recognition using a machine-learning model. For the categorization of with mask and without mask on our dataset, we have showed a considerable improvement in performance over publicly maintained datasets and open source tools. Our designed model offers 99% accuracy. Despite the medium size of the publicly accessible dataset, the results are encouraging. This use of face mask detection and an alertsystem is more beneficial for preventing the spread of COVID and locating and identifying those who disregard safety precautions in public places. We can train the model on more data in future work to acquire even better results and to assess more face mask detection architectures. This technology can be uploaded to the cloud to rapidly offer detection results and to assist the law enforcement by locating and identifying those who are not wearing masks. The workload of the authorities' staff should drastically decrease as a result. Also, we can use this technology in public areas like ATMs, banks, and so forth.

## References


1. *Mrs. Monali Nitin chaudhari, Ms. Mrinal Deshmukh, Ms. Gayatri Ramrakhiani, Ms. RakshitaParvatikar, ''Face detection using viola-jones algorithm and neural networks'', IEEE, 978- 1-5386-5257-2/18, 2018.*
2. *Jie Wang, Zhao Li, '' Research on face recognition based on CNN'', 2nd International Symposium on Resource Exploration and Environmental Science IOP Conf. Ser.: Earth Environ. Sci. 170 032110, 2018.*
3. *SukhadaChokkadi, Sannidhan MS, Sudeepa K B, AbhirBhandary, ''A study on various state of the art of the art face recognition system using deep learning techniques'', International Journal of Advanced Trends in Computer Science and Engineering, Vol 8(4), July- August 2019.*
4. *Bosheng Qin and Dongxiao Li, '' Identifying facemask-wearing condition using image superresolution with classification network to prevent covid-19'', Sensors, Vol. 20, 2020.*
5. *Vinitha.V and Velantina.V'Covid-19 facemask detection with deep learning and computer vision'', International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 08, Aug 2020.*
6. *Yan, K., Huang, S., Song, Y., Liu, W., & Fan, N. (2017). Face recognition based on convolution neural network. 2017 36th Chinese Control Conference(CCC). doi:10.23919/chicc.2017.8027997*
7. *Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement, 167, 108288.*
8. *Ge, S., Li, J., Ye, Q., & Luo, Z. (2017). Detecting Masked Faces in the Wild with LLECNNs. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 426–434*
9. *.Xin Zhang, Thomas Gonnot, JafarSaniie, "Real-Time Face Detection and Recognition in Complex Background", Journal of Signal and Information Processing, 2017, 8, 99-112.*
10. *K. Sunil Manohar Reddy, "Comparison of Various Face Recognition Algorithms", International Journal of Advanced Research in Science, Engineering and Technology Vol. 4, Issue 2, February 2017.*
11. *S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, "Rational use of face masks in the covid19pandemic,"The Lancet Respiratory Medicine, 2020.*
12. *Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Peiet al., "Masked facerecognition dataset and application,"arXiv preprint arXiv:2003.09093, 2020.*
13. *Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review,"IEEE transactions on neural networks and learning systems, vol. 30, no. 11, pp. 3212–3232, 2019.*
14. *A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review, "Artificial Intelligence Review, vol. 52,no. 2, pp. 927–948, 2019.*
15. *D.-H. Lee, K.-L. Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu, "Deep learning and control algorithms of direct perception for autonomous driving,2019.*