International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Real Time Face Mask Detection using Deep Learning Approach

[1]Shikha Singh,[2]Vandana Dubey,[3]Shalini Agarwal,[4]Brahma Hazela,[5]Sheenu Rizvi,[6]Vineet Singh

[1,3,4,5,6] Amity School of Engineering and Technology, Amity University, Uttar Pradesh, Lucknow Campus, INDIA
[2]C. Byregowda Institute of Technology, Kolar, Karnataka, INDIA

Abstract— In response to the need for face masks in preventing dangerous illnesses like COVID-19, this paper is centered around deep learning-based face mask identification. Leveraging powerful libraries such as TensorFlow, Keras, and OpenCV, our objective is to create a reliable and effective system capable of discerning individuals wearing face masks from those who are not. This innovative approach holds immense potential for applications in public health monitoring, security systems, and critical public spaces where adherence to face mask policies is of utmost importance.To ensure practical usability, we deploy the trained model in real-time scenarios, utilizing webcams or video streams. The system efficiently processes frames, swiftly detecting and classifying faces with or without masks, and promptly providing feedback or alerts to users or surveillance systems.This paper highlights how well TensorFlow, Keras, and OpenCV work together to create a reliable and accurate face mask detection system. The amalgamation of deep learning, image processing, and real-time capabilities facilitates seamless monitoring of face mask adherence, significantly contributing to public safety and health initiatives across various domains.

Keywords- Mask Detection, Connvolutional Neural Network, Keras, OpenCV, Tensorflow

I. INTRODUCTION

According to the latest official Situation Report (205) by the United Nation's World Health Organization (WHO), if we seek out from the advent of the global pandemic then the virus has affected over a million of individuals and this population can account to the living population of one whole nation [1]. It is important for the public to understand when and how to wear masks for source control and prevention of COVID-19. Wearing masks can reduce the risk of exposure to harmful particles emitted by an infected individual, even during the "pre-symptomatic" phase. Since wearing a mask effectively stops the infection from spreading, it also aids in destigmatizing the practice. In order to manage the pandemic, the WHO underlines the use of lifesaving masks and breathing devices for health care providers. Consequently, in the present-day global culture, it has become crucial to get face masks. Contagious illnesses like COVID-19 have recently spread quickly around the world, causing severe disruption and posing serious hazards to the public's health. The application of face masks that are effective is a vital preventative practice that has earned global attention for helping to stop the global propagation of such infections. By limiting the dispersion of droplets generated by breathing, mask use in open areas lowers the risk of viral transmission. The necessity for effective and precise identification of face mask devices has grown critical as the demand for face masks rises. Older manual inspection procedures take a lot of time, are prone to mistakes, and cannot handle big populations. Because of this, there is an increasing interest in using machine learning approaches to automate the face mask identification process [2]. With this paper, we intend to make use of neural networks to construct an automated facial mask recognition system, particularly by combining TensorFlow, Keras, and OpenCV.

The paper is set up as follows. The Introduction is Section 1, and the Literature Review, which summarizes and provides the results of earlier work that served as the background research for this paper, is Section 2. The Methodology used for this project and a flowchart of the suggested procedure are both found in Section 3. The Experimental Analysis in Section 4 , the Result and discussion is highlighted in Section 5 and Section 6 is where this paper reaches its Conclusion.

II. LITERATURE REVIEW

In this study, an extensive review of the existing literature has been performed. Based on the findings from numerous articles, the primary focus of research has largely been on enhancing prediction accuracy through the utilization of machine learning techniques, yielding positive results.

The researchers utilized both single and multi-stage detection techniques, including CNN, R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN, and YOLO. Their proposed model achieved a recall rate of 96.69% and a runtime of 130ms per frame. Among the various methods tested, Faster R-CNN exhibited the shortest running time, while the suggested approach performed comparably to Mask R-CNN in terms of elapsed time [3].

The researchers employed a two-stage detection approach to detect faces and apply the model. The R-CNN method employed a peculiar recognition technique to firstly derive an inventory of object, followed by using an SVM classifier to predict objects and their corresponding classes. SPPNet, on the other hand, modified R-CNN by adding an SPP layer and using features from several area suggestions that were then put through an entirely interconnected layer for classification. The model's reliability in identifying whether or not someone is concealing themselves with a mask reached up to 95% [4].

The Residual Neural Network (ResNet 50) technique was used by the researchers to extract characteristics from the photos. Several typical machine learning algorithms—Support Vector Machine (SVM), Decision Tree (DT), and Ensemble Learning (EL)—were subsequently trained using the attributes that had been retrieved. The classifiers were tested for facial mask spotting after development. The SVM classifier has the best detection accuracies out of all the classifiers. The SVM classifier attained detection accuracies of 99.64% and 99.49% in the RMFD and SMFD datasets, respectively. Additionally, the SVM classifier had a flawless detection accuracy of 100% in the LFW dataset [5].

The authors trained a model especially intended for facial mask identification using deep learning techniques. A sizable dataset of pictures featuring individuals donning face shields and pictures of folks lacking masks was used to train the algorithm. The simulator successfully acquired the ability to discriminate between uncovered and concealed faces during this training method. The authors ran tests with data from the actual world to assess the effectiveness of their method. The technology was put to the test using pictures and video streams taken in a variety of settings, such busy venues or subway stations. The results of this particular study revealed the automated system's accuracy rate for spotting face masks. The algorithm was quite accurate in determining whether people were wearing masks or not. The authors reported the performance metrics achieved by their system, to assess its effectiveness comprehensively [6].

The scholars proposed a model using a DL approach to ascertain whether someone is using a mask while in public. For this, the authors used the MobileNetV2 picture categorization method. They used two datasets: one from Kaggle called RMFD, while the other was compiled from surveillance footage, road signal cameras, and retail cameras in 25 cities throughout Indonesia. The model was trained using these datasets. On the experimental subsets of the RMFD and Indonesian datasets, the trained model demonstrated detection accuracies of 95 percent and 85 percent or higher, respectively [7].

More than one DL approaches, namely DCNN and companying it was the MobileNetV2 are suggested, for our evaluation. To assess the performance of these models, we have utilized two distinct datasets. Dataset-1 comprises 2500 images with and without masks, while dataset-2 consists of 4000 images with and without masks. For training the MobileNetV2 architecture, we have employed 80% of the data from each dataset, reserving the remaining 20% for model testing. As for the DCNN model, 90% of the data from each dataset has been utilized for training, with the remaining 10% set aside for testing. The MobileNetV2 model achieved impressive training and validation accuracies on both datasets. For dataset-1, the training accuracy reached 99%, while the validation accuracy attained 98%. Similarly, for dataset-2, the training accuracy was 99% and the validation accuracy was 99%. In the case of the DCNN model, it achieved solid training and validation accuracies as well. The training accuracy on the two data sets was 98%, whereas the validation accuracy was 97%. We have achieved excellent accuracy rates by applying these DL frameworks and carrying out extensive training and assessment on the datasets, demonstrating the usefulness of the models in determining the fact that people are sporting masks [8].

A technique that makes use of the MobileNetV2 architecture to find masks across both footage and picture sources. When tested against a set over four thousand or so photographs, the system had an impressive identification rate of 98%. The authors emphasize that their approach can analyse real-time visual data and is appropriate for devices with constrained computational capabilities [9].

Convolutional Neural Network (CNN) layers serve as the primary architecture in the model proposed in [10] for the construction of various layers. The model is divided into three main sections: face mask detector application, CNN model training, and data pre-processing. Roughly 4,000 photos from a dataset gathered from Kaggle are used for the model's training and assessment. A superior reliability rating of 98% is given by the model. In comparison to current algorithms such as DenseNet-121, MobileNet-V2, VGG-19, and Inception-V3, the suggested model displays more algorithmic precision as well as effectiveness.

A GAN-based algorithm was created by the team of investigators in [11] to handle the demolition of recognized masks from the face, the creation of absent cosmetic aspects with better details, and the rebuilding of areas. Their suggested GAN architecture included separate discriminators: one was designed to capture the overall shape of the mask worn on the face, and the other was intended to remove the area that the cover had occluded. Two simulated data sets were used in the model building procedure.

A mobile phone-based detection approach was put out by the researchers in [12]. From gray-level co-occurrence matrices (GLCMs) of micro-images of face masks, they derived three elements. A three-result recognition study using the K-nearest neighbors (KNN) method was completed with a total recognition precision of 82.87%. The grey-level co-occurrence vector was used by the algorithm to find masks with faces in microphotos. It is crucial to remember that this style was created especially for handheld devices and might not be appropriate for applications other than smartphones.

[13] Introduces a personalized support vector machine (SVM) driven real-time pose detection and correction system. This system recommends yoga practices tailored to individual health conditions or diseases, aiming to offer a dependable and easily accessible resource for individuals seeking to utilize yoga as a complementary approach in managing their health conditions.

[14] study proposes an automated technique for human face identification and recognition, leveraging facial characteristics extracted through deep learning. The high accuracy rate of deep learning makes it a suitable method for conducting face recognition tasks.

III. METHODOLOGY

To develop a custom face mask detector, our project can be divided into two main phases, each consisting of specific sub-steps. The above given Fig.1 is the graphical representation of the steps and phases this project is broken down into:

Training: During this phase, our primary focus is on preparing the face mask detection dataset, training a model utilising Keras and TensorFlow, and preserving the trained face mask detector for usage in subsequent phases.

Deployment: Once the face mask detector has been trained, we will be able to move on to the deployment part of the process. To accomplish this, load the trained model, detect faces in photos or video streams, and then categorize each face observed as "with_mask" or "without_mask.".

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

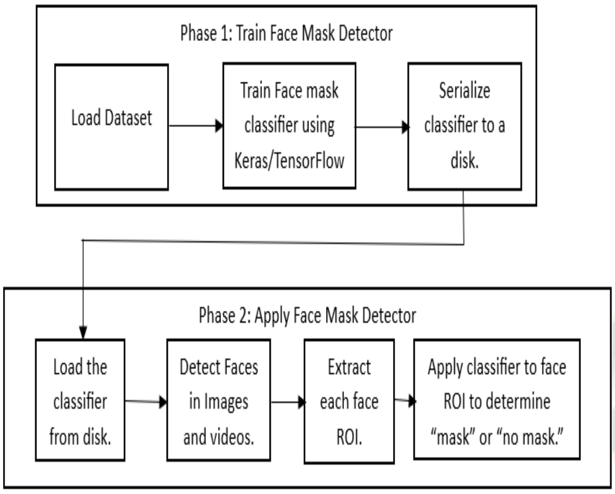

Fig. 1. Flowchart for the steps involved in the project.

A. *Dataset Used*

Face mask identification was done using the Kaggle dataset, which is a well-liked dataset for developing and testing face mask recognition models. Pictures that fall into the "with_mask" and "without_mask" categories make up the dataset. It comprises 7553 RGB images that have been separated into two files based on whether or not masks are included. There are specifically 3828 photographs of faces without masks and 3725 photos of faces with masks.A total of 1776 photographs representing both "with_mask" and "without_mask" samples were obtained from Prajna Bhandary's GitHub account in order to increase the dataset. In addition, 5777 more photos from Google searches were gathered and reviewed for appropriateness and relevancy. The dataset contains a wide variety of images that were taken in a variety of settings, including both indoor and outdoor settings, varying lighting situations, and people of all ages, genders, and nationalities. The goal is to offer a sample that is representative of the real-world situations in which face mask detection systems are anticipated to function [15]. Below given figures Fig.2. and Fig.3. shows a few samples of data set inputs that are fed to the classifier.

Fig. 2. "Without_mask" dataset

Fig. 3. "With_mask" dataset

B. *Preprocessing Data*

The following steps are typically performed as part of the data preprocessing phase: Loading and preprocessing the dataset: The images from the dataset are loaded and processed. Each image is resized to (224, 224) pixels, converted to an array, and preprocessed using the MobileNetV2 preprocessing function. Performing label encoding: The labels (with_mask and without_mask) associated with each image are encoded using one-hot encoding. Splitting the dataset: This ensures that the model is evaluated on unseen data during testing.

C. *Loading the Base Model*

By loading the MobileNetV2 model as the base model and excluding the head layer, we can leverage the pre-trained features learned by the MobileNetV2 model for our face mask detection task. This allows us to benefit from the model's ability to extract meaningful features from images, which can be useful in detecting whether an individual has covered his face using a surgical mask or his face is uncovered [16]. An AveragePooling2D layer is added to perform average pooling on the output tensor. Average pooling decrements the attribute of dimensions related to that of space of the tensor while not tampering with the said important features. Here, a pooling window of size 7x7 is used, which means that the tensor will be divided into non-overlapping windows of size 7x7, and the average value within each window will be computed. The pooled tensor is flattened into a 1-dimensional vector using the Flatten layer. This gets the data ready for the fully connected layers that come after. 128 neurons make up the completely connected Dense layer that is added. The network is given nonlinearity by using the relu activation function. This layer develops a higher-level representational mapping for the flattened features. Inserted is a dropout layer with a dropout rate of 0.5. Dropout introduces regularization during training by setting a portion of the input units to 0 at random, hence assisting in preventing overfitting. Two neurons are introduced to the last dense layer. By generating probabilities for the two classes, "with_mask" and "without_mask," this layer completes the classification operation. The projected probability must add up to 1 using the softmax activation function. By constructing the top spot of the model and connecting it to the output of the base model, we create a complete model for face mask detection. The base model leverages pre-trained weights for feature extraction, while the head model learns to classify the extracted features.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

D. Training the Model

The model is trained using an annotated sample of images containing and excluding face masks. While the images are being fed through the network, the network parameters are updated using optimization techniques such stochastic gradient descent (SGD) or Adam, which computes the disparity among what is expected and the surface truth output. The model can be used to forecast new images by feeding new ones through the network and gathering the estimated probabilities after training. The predicted class is the one that has the best chance of occurring. The model.fit() function is used for the training phase, which maximizes the weights for the model based on the provided training data. On combining the training as well as the validation datasets, the model makes modifications to the parameters during training to decrease loss and increase accuracy. Predictions are made after model training is finished.

E. Model Evaluation

After being done with model training we evaluate the model accuracy. Various performance metrics are employed to analyse the model accuracy. The classification_report() function is used to generate a classification report based on the true labels (testY.argmax(axis=1)) and the predicted labels (predIdxs). The testY.argmax(axis=1) expression converts the one-hot encoded true labels (testY) into their corresponding class indices. This is necessary to match the format of the predicted labels (predIdxs).
The model accuracy came out to be as around 99% which can be seen in Table 1.

TABLE I. TABLE SHOWING MODEL ACCURACY BASED ON MULTIPLE PERFORMANCE METRICS.

	Precision	*Recall*	*F1-score*	*Support*
with mask	0.99	0.99	0.99	383
Without mask	0.99	0.99	0.99	384
Accuracy			0.99	767
Macro avg	0.99	0.99	0.99	767
Weighted avg	0.99	0.99	0.99	767

F. Extracting Face ROI

Extracting face ROI refers to the process of isolating the region within an image or frame that contains a person's face. The face ROI is typically defined by a bounding box that tightly encloses the facial features of an individual. [17]
The confidence (probability) corresponding to each detection is extracted. Filtered out are weak detections that have a confidence level below a predetermined threshold (in this case, 0.5). Based on the location and size of the detection, the bounding box's (x, y)-coordinates are calculated and then modified to fit into the frame's boundaries. The determined the box boundaries are used to extract the face region from the frame. The face ROI's color space is changed from BGR to RGB. The face ROI is shrunk to (224, 224) pixels in fixed size. Using img_to_array, the facial ROI is transformed into a NumPy array. The pre-processing of the face ROI often includes standardizing the values of each pixel. The bounding box coordinates for the beforehand processed face ROI have been included to the existing locs list, and the faces list is updated with their addition. The facial regions that are important are retrieved, processed, and ready for categorization using the facial masked identification model thanks to this procedure.

G. Real Time Predictions

The frame is resized to have a maximum width of 400 pixels using imutils.resize. The detect_and_predict_mask function is called to detect faces and classify whether they are wearing a mask or not. The function returns the face locations (locs) and their corresponding predictions (preds). A loop is used to iterate over each detected face and its prediction. The bounding box coordinates and predictions are unpacked. Based on the prediction, the class label ("Mask" or "No Mask") and colour (green for "Mask" and red for "No Mask") are determined. The label, including the probability, is displayed above the bounding box using cv2.putText. The bounding box rectangle is drawn around the face using cv2.rectangle. The output frame with labeled faces is displayed using cv2.imshow. The loop continues until the "q" key is pressed, at which point the loop is broken and the program terminates.

IV. EXPERIMENTAL ANALYSIS

According to Fig. 4, which compares the training and validation loss for the dataset in table 1, the model achieves a precision of 99%. The use of MaxPooling is one of the main elements causing this great accuracy. This procedure decreases the amount of parameters that need to be learned while introducing fundamental translation invariance to the model's internal representation. The dimensionality is successfully decreased by discretizing the input representation, which in this example is a picture. The model's optimal number of neurons is 64, which strikes a compromise between complexity and efficiency. Performance might suffer if neurons and filters are set to a significantly larger number. The crucial area of the image, the face, is efficiently recorded for precise mask identification without the risk of overfitting because to the optimum values for the filters and pool_size. Abbreviations and Acronyms

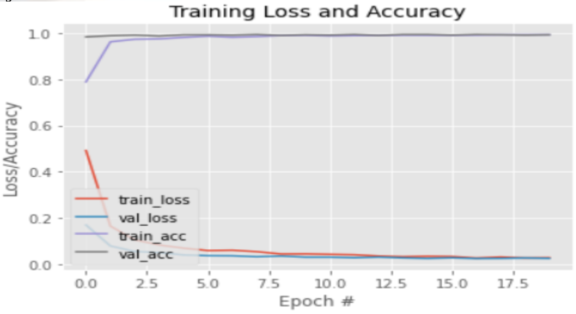

Fig. 5. Graph showing model loss and accuracy.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

V. RESULTS AND DISCUSSION

The face detection model saved is accessed through the Anaconda prompt terminal. There are two python files one is for training the model over the static data set provided as input and the other one is for loading that trained and saved model and preparing it to perform video stream or real scenario predictions.

Now we can access the python file through the prompt and perform our face mask detections. The model is capable of detecting multiple pupils at a single instance.

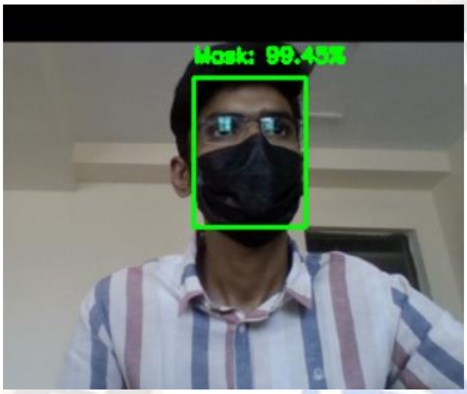

Fig. 6. Detecting the face of individual wearing mask in real time.

It is obvious in the above Fig. 6. that whenever a protective mask is worn appropriately, concealing one's mouth and nostrils, a small green box with a sealed surround appears around the wearer's face on the detector.

Fig. 7. Detecting the face of an individual not wearing a mask in real time.

As we can see in the above Fig. 7 that if an individual is not wearing a mask or is wearing one but not placing it in the manner deemed appropriate covering the mouth and nostrils then the box that appears on the screen is red in colour instead of green and it is the indication that the person needs to wear a mask correctly.

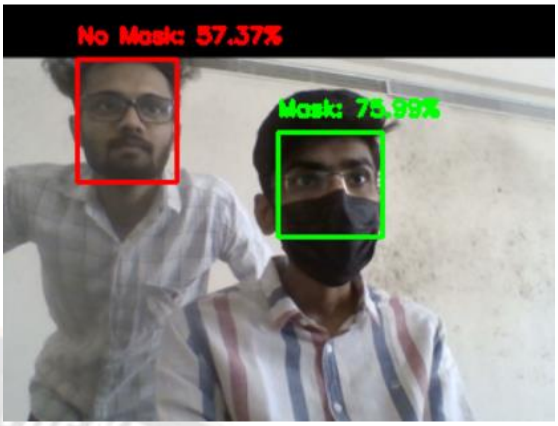

Fig. 8. Detecting multiple individuals together in real time.

VI. CONCLUSION AND FUTURE SCOPE

Now to sum it all up regarding this project, we utilised Keras, OpenCV, and TensorFlow, the face mask identification project has been a stunning success, obtaining an exceptional accuracy of 99%. This project is highly practical and relevant in a variety of settings because it not only emphasizes accuracy but also shows how to carry out real-time detections on video streams. We have developed a reliable and effective system that can correctly recognize people wearing face masks in real time by utilizing the potent combination of Keras, OpenCV, and TensorFlow. A complete framework for preprocessing photographs, deep learning model construction, and real-time video stream processing has been made possible by the integration of Keras, OpenCV, and TensorFlow. In contrast to OpenCV, which provides a broad variety of image processing and computer vision features, Keras makes the process of creating and training deep learning models simpler. TensorFlow acts as the foundation, offering the required infrastructure for effective computation and deep learning model optimization. The choice of deep learning architecture, which uses convolutional neural networks (CNNs) to extract pertinent characteristics and provide precise predictions, is beneficial as evidenced by the accomplishment of 99% accuracy. As for the scope of the future with respect to work in the field of face mask detection one can say, in recent years, face mask detection has come to assume an increasingly vital role as a direct result of the COVID-19 pandemic.

.

REFERENCES

[1] Adusumalli, H., Reddy, S. V., Akula, S. P., & Srikanth, M. "Face mask detection using OpenCV." In: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). IEEE, 2021.

[2] Shete, S., Tingre, K., Panchal, A., Tapse, V., & Vyas, B. "Mask Detection and Tracing System." (2021).

[3] Ali, M., Amer, F., & Al-Tamimi, M. S. H. "Face mask detection methods and techniques: A review." International Journal of Nonlinear Analysis and Applications, 13(1), 3811-3823.

[4] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic."(2021)

[5] Militante, S. V., & Dionisio, N. V. "Realtime facemask recognition with alarm system using deep learning." In: 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), pp. 106-110, IEEE, Shah Alam, Malaysia, 2020.

[6] Suresh, K., Palangappa, M. B., & Bhuvan, S. "Face mask detection by using optimistic convolutional neural network." In: 2021 6th International Conference on Inventive Computation Technologies (ICICT), IEEE, 2021.

[7] Rahman, M. M., Manik, M. M. H., Islam, M. M., Mahmud, S., & Kim, J. H. "An automated system to limit COVID-19 using facial mask detection in smart city network." In: 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Vancouver, BC, Canada, 2020.

[8] Sanjaya, S. A., & Rakhmawan, S. A. "Face mask detection using MobileNetV2 in the era of COVID-19 pandemic." In: 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), pp. 1-5, IEEE, Sakheer, Bahrain, 2020.

[9] Oumina, A., El Makhfi, N., & Hamdi, M. "Control the COVID-19 pandemic: face mask detection using transfer learning." In: 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, 2020.

[10] Hussain, D., Dostdar, R., Siddiqui, S., & Saeed, A. "Face mask detection using deep convolutional neural network and MobileNetV2-based transfer learning." Wireless Communications and Mobile Computing, 2022.

[11] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.C. "Mobilenetv2: inverted residuals and linear bottlenecks." In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510-4520, Utah, 2018.

[12] Goyal, H., Sidana, K., Singh, C., Jain, A., & Jindal, S. "A real time face mask detection system using convolutional neural network." Multimedia Tools and Applications, 81(11), 14999-15015, 2022.

[13] Mane, D. ., Upadhye, G. ., Gite, V. ., Sarwade, G. ., Kamble, G. ., & Pawar, A. . (2023). Smart Yoga Assistant: SVM-based Real-time Pose Detection and Correction System. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7s), 251–262. https://doi.org/10.17762/ijritcc.v11i7s.6997

[14] Jatain, R. ., & Jailia, M. . (2023). Automatic Human Face Detection and Recognition Based On Facial Features Using Deep Learning Approach. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 268–277. https://doi.org/10.17762/ijritcc.v11i2s.6146

[15] Din, N. U., Javed, K., Bae, S., & Yi, J. "A novel GAN-based network for unmasking of masked face." IEEE Access, 8, 44276-44287, 2020.

[16] Chen, Y., Wang, J., Chen, Z., & Yang, Z. "Face mask assistant: Detection of face mask service stage based on mobile phone." IEEE Sensors Journal, 21(9), 11084-11093, 2021.

[17] Shah, R. C., & Shah, R. J. "Detection of face mask using convolutional neural network." Mobile Information Systems, 43, 382-487, 2019.