Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A novel machine learning scheme for face mask detection using pretrained convolutional neural network

T.M. Saravanan*, K. Karthiha, R. Kavinkumar, S. Gokul, Jay Prakash Mishra

Department of Computer Applications, Kongu Engineering College, Perundurai 638 060, Tamil Nadu, India

A R T I C L E   I N F O

Article history:
Available online 21 January 2022

Keywords:
Face mask detection
Vgg16
Transfer learning
Image augmentation and machine learning scheme

A B S T R A C T

Corona virus 2019 (COVID-19) erupted toward the end of 2019, and it has continued to be a source of concern for a large number of people and organizations well into 2020. Wearing a face cover has been shown in studies to reduce the risk of viral transmission while also providing a sense of security. Be that as it may, it isn’t attainable to physically follow the execution of this strategy. This proposed system is built by pretrained deep learning model, Vgg16. The proposed scheme is easy to implement and use all the layers in vgg16 model and train only the last layer called fully connected layer, which reduce the training time and effort. The proposed scheme is trained and evaluated using two Face mask datasets, one having 1484 pictures and the other with 7200. For a smaller dataset, augmented pictures were utilized to enhance accuracy. The suggested model is tested on unknown pictures, and it correctly predicts whether the image is wearing a mask or not. The proposed scheme gives accuracy 96.50% during testing in small dataset. The model gives accuracy in medium dataset is 91% during testing. By using vgg16 pretrained model and image augmentation in the dataset improves performance and gives a high accuracy.

1. Introduction

The COVID-19 corona virus pandemic is wreaking havoc on the world’s health. Using a mask in public places and in congested regions is the most effective COVID-19 prevention strategy. In these places, personally monitoring individuals is quite tough. For face mask identification, a hybrid model combining deep and conventional machine learning will be presented [1]. Since the outbreak of the Covid-19 epidemic, substantial progress in the fields of image processing and computer vision has been made in the identification of face masks. Several methods and strategies have been used to construct several face detection models. Due to the unexpected appearance of the COVID-19 epidemic, different facial recognition technologies is currently being used on persons wearing masks. Confront detectors face a difficult job in detecting face masks [2] (see Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12).

Corona virus (COVID-19) is a corona virus family member. This epidemic brings the entire globe to a halt and triggers a worldwide economic downturn. Face recognition software is used to manage complicated images. To extract the face region from the full image, the face detection method is used first. Face detection is carried out using a color-based algorithm. After face localization, a face mask detection technique is used to determine whether or not a face is wearing a mask [3]. The challenge of identifying whether a person is wearing a face mask from speech is useful in forensic investigations, surgeon communication, and persons protecting themselves against contagious illnesses like COVID-19. For mask identification from speech, we offer a unique data augmentation approach [4] (see Tables 1 and 2).

Face Mask Detection Platform recognizes whether or not a user is utilizing a machine learning or deep learning scheme to wear a mask. Face masks have replaced conventional methods of protection during the pandemic, as they are efficient in stopping viral transmission. Several industrialized and developing countries across the globe have made it necessary for people to wear masks while they are at home or in public areas.

Because of the fast spread of corona virus, the globe is facing a major health disaster COVID-19. Using a mask in public places and in congested regions is the most effective COVID-19 prevention strategy. Because personally monitoring individuals in these places is extremely difficult, Face Mask Detection serves a critical func-
tion. A transfer learning scheme is provided to computerize the process of detecting people who are not wearing masks.

2. Literature review

Nanette et al. (2020), the Covid-19 commonly known as serious acute respiratory disorder, is a disease that causes serious respiratory problems. This disease is an infectious illness spread by respiratory droplets from an infected unwell person who talks, sneezes and coughs. This spreads fast all the way through close contact with infected people or by touching contaminated goods or surfaces. There is presently no vaccination available to defend against Covid-19; therefore avoiding infection appears to be the merely way to guard ourselves. In public, exhausting a facemask conceals the nose and mouth. As technology has evolved, deep learning has proven its usefulness in image processing detection and classification. Deep learning approaches for facial recognition and determining whether or not a person is putting on a facemask. The dataset gathered has a 96% accuracy rate when it comes to the trained model's performance. If the individual spotted is not wearing a facemask, the system creates a raspberry Pi-related real-time facemask identification system which alerts and records facial image [5].

Anushka et al. (2020), the corona virus outbreak has had a devastating impact on the entire planet. According to the World Health Organization, wearing a mask is now required to prevent the transmission of the infection, among other things. To prevent acquiring the fatal illness, everyone in the country prefers to live a healthy lifestyle by wearing a mask in public meetings. Recognizing faces wearing masks is a difficult task since there are few datasets available that include both masked and unmasked pictures. A layered Conv2D model for detecting facial masks that is quite effective. We used Gradient Descent for training and binary cross entropy as a thrashing function to build this scheme, which is a stack of 2-D convolutional layers with RELU activations and Max Pooling. The model was trained using a combination of two datasets. A 95% validation/testing accuracy were achieved [6].

Jiang and Fan (2020), corona virus illness has had a major impact on the world in 2019. Wearing masks in public places is one of the most common ways for individuals to protect themselves. Many public service providers only allow clients to utilize their services provided they appropriately wear masks. Nevertheless, there are just a few research works on image analysis-based face mask identification. Retina Face Mask is a face mask detector with great accuracy and efficiency. Retina Face Mask is a one phase detector that uses a attribute pyramid network to blend high level semantic data with numerous feature maps, as well as an unique context attention module to identify face masks. To decline predictions with little confidence and a high intersection of union, use a cross-class object removal method [7].

Chen et al. (2020), since its huge outbreak during December 2019, the Covid-19 have increased all over the globe, causing a tremendous loss to the entire world. Wearing masks is a simple and efficient way to stop it from spreading at the source. Face masks are typically used infrequently but for a short period of time. We propose a detection technique based on the mobile phone to address the problem of not knowing which service stage of the mask belongs to. Four characteristics may be extracted from the GLCMs of the face mask micro-photos. The KNN algorithm is then used to create a three-result detection system. Validation studies reveal that the system can achieve an accuracy of 82.87 percent on the testing dataset [8].

Ejaz et al. (2019), principal component analysis was applied to mask and non-mask face identification. In today’s world, the phrase “security” is crucial. Face recognition is extensively used in biometric technology to protect any system since it is superior to other conventional approaches such as PIN, password, fingerprint, and so on, and it is the most dependable way to identify or verify a person effectively. These kinds of masks have an impact on the accuracy of facial recognition. Many non-masked facial recognition algorithms have recently been created that are extensively utilized and provide higher performance. PCA is a more successful and effective statistical approach that is extensively used [9].

Meenpal et al. (2019), in image processing and computer vision, face detection has become a highly popular problem. Convolutional architectures are being used in many new algorithms to make them as accurate as feasible. These convolutional designs have allowed even pixel information to be extracted. The method for creating accurate face segmentation masks from any image of any size. Fully Convolutional Networks are used in the training to semantically separate out the faces in the image. The FCN’s output picture is cleared up to eliminate unnecessary noise and prevent erroneous predictions [10].

Zhang et al. (2019), Face identification and alignment in an unrestricted environment are difficult problems to solve, yet they are necessary for preserving traffic tidy and public security. The enhanced multi-task cascaded convolutional networks to achieve finest face area recognition and attribute alignment of the driver face on the road, and to forecast face and attribute placement using a coarse to fine pattern. In addition, ITS-MTCNN approach proposes an enhanced regularization method as well as an effective online hard sample mining methodology. In the self built driver face database, the training scheme and divergence experiment are run. Finally, comparison studies display the efficacy of the ITS-MTCNN scheme [11].

Wu et al. (2019), for accurate face identification, significant representations for describing face appearance were required. Current detectors, especially based on convolutional neural networks applies functions like convolution to any native areas on every face for attribute aggregation and consider all native options to be uniformly effectual for the recognition task. Specified a proposal where part-specific attentions sculptural by learnable gaussian kernel that will be used to look for correct native area placements and scales in order to dig out dependable and revealing facial element alternatives. Later, using LSTM, face specific notice is used to model relationships amid native components and modify their contributions to detection tasks. To illustrate the efficiency of our hierarchic attention and construct comparisons with progressive methods, we conducted in-depth tests on three tough face detection datasets.

Vallimeena et al. (2019), CNN algorithms have been increasingly popular in recent years for a variety of computer vision-based applications, such as disaster management systems that use crowd-sourced pictures. Flooding is a common natural catastrophe that poses a hazard to human lives and property. Flood photos with individuals recorded by smart phone cameras are being used to calculate the depth of the water to determine the amount of damage in flood-affected areas. For these purposes, a variety of CNN algorithms are available. Each one has a different architecture, which has an impact on the accuracy of the results [13].

Set et al. (2019), Deep Metric Learning is used to detect and identify human faces. The approach employed in this work is called deep metric learning, and it combines face identification with human object detection. Real-time pictures or stream movies collected by CCTV or any other video capturing equipment may be used to analyse captured footage. Our approach uses widely known classification algorithms to recognise and identify faces even in blurry pictures or screen-together movies [14].

Kumar et al. (2019), face recognition for diverse applications rely heavily on multiple face detection and extraction. Support vec-
tor machine was utilized for multiple face recognition in the suggested approach, while Discrete Wavelet Transform, Edge Histogram, and Auto-correlogram were used for extracting features. For MFD, suggested technique was tested on two distinct databases: Carnegie Mellon University and BAO. The anticipated scheme performs superior than the existing strategy in this study article. Finally, our accuracy increased to around 90% [15].

3. Related work

The improved viola-jones face detection scheme that is based on Holo Lens [16], which improves the classic viola-jones face detection scheme by relying on Haar like rectangle attribute expansion to improve detection competence and accelerating recognition building using 2D convolution separation and image Resampling techniques. Detecting face masks has become a critical duty in aiding worldwide civilization. The method properly recognizes a face from a picture and then determines if it has a mask on it or not, and it can also detect a face and a mask in motion. To accurately identify the existence of masks without generating
over-fitting, a Sequential Convolutional Neural Network model was used [17]. Using the Haar Cascade approach, data extraction and human face detection may be done. The Haar Cascade technique may be used to filter selfie face pictures with a high level of accuracy [19]. Previous models were primarily trained and evaluated on high-quality pictures, which isn’t necessarily the case in real-world applications like surveillance systems. While testing on low-quality pictures of various levels, performance degraded [20]. For the opinion-examination of restaurant evaluations and sentence-based opinion summary analysis, they introduced the traditional and VADER opinion mining schemes [20].

4. Proposed scheme

Many of the articles in the aforementioned literature discovered that deep learning and machine learning techniques were utilized to classify whether or not a person was wearing a mask. They train a model from scratch that is computationally costly and prolonged. On real world pictures, the performance fails. In this article, data augmentation and a pre-trained convolutional neural network scheme named vgg16 are used to extract features and categorize the picture as with or without mask to obtain greater accuracy. This improves the dataset’s accuracy while also being simple to apply. The proposed method overcomes the disadvantage of the existing method.
5. Research methodology

There are several phases involved in detecting face masks, and below is a graphical depiction of the scheme.

5.1. Data collection

The process of gathering data from kaggle is known as data collection. For Face mask identification, we collected two datasets from the Kaggle website, each containing training, validation, and testing pictures.

5.2. Image resizing

The collection comprises pictures of various sizes. The first step is to resize the image to 224 * 224 * 3, where 3 denote RGB color (red, green, blue). Before sending a picture to the model, every image in the dataset is scaled to 224 * 224 pixels.

5.3. Image augmentation

It is an approach for artificially increasing the range of a training dataset by manipulating pictures within it. The training pictures of a smaller dataset are enhanced in this article, including rescaling, width, height, horizontal flipping, rotating and zooming.

5.4. Pretrained convolutional neural network

In image categorization, a convolutional neural network i.e. Deep Neural Network is utilized. CNN image classification takes an input image and processes it before categorizing it (Wearing Mask or Not Wearing Mask).

To categorize pictures with mask and without mask, this study employs one of the pre-trained models — VGG 16 with Deep Convolutional Neural Network. To train quicker and more cost-effectively, the vgg16 pre-trained model, also known as transfer learning, was developed.

Convolution layer is used in the Vgg16 model to extract the feature map without changing the size of the input and the output of

![Fig. 11. Training and validation loss for Dataset 2.](image)

![Fig. 12. Output of given input image.](image)
the convolution level is sent to the max pooling level. By extracting just feature from the input, the yield of the convolution layer is provided as input to the Max pooling level, which reduces the picture by half. The vgg16 model extracts features using 13 convolution layers and five max pooling layers. These two layers take features from the input and provide a feature vector as an output. Pass the output to the fully linked layer first that is converted from a feature vector to a 1D array. An output layer that is fully linked is utilized to categorize the pictures. Soft Max is an activation function that is used to determine which class has the greatest likelihood. The Soft Max function is used to classify binary and multi-class data. In this suggested technique, all layers are frozen, and we don’t need to train the entire network for our dataset; instead, we may just utilize the data from the prior model. We’re not going to use the last layer since we’re weary with the pre-created network; instead, we’ll use our own custom FCL to train our dataset. We’ll add one output layer on top of FCL and then integrate our fully connected stratum with the pre trained network.

The Phrase transfer learning refers to describe a scheme that has been previously transferred. Transfer learning is a technique for applying previously learned model information to a new task. To categorize pictures as with mask or without mask, this project employs one of the pre-trained models VGG 16 with Deep Convolutional Neural Network. Transfer learning saves a significant amount of computational power. All vgg16 levels, excluding completely linked layers, are frozen in the suggested technique. Only the last layer should be retrained.

### 6. Result and discussion

Image augmentation and resizing are performed on two distinct datasets i.e. Face mask dataset and dataset train on the pre trained vgg16 model, also known as transfer learning, which yielded excellent accuracy and required minimal computing time.

Face masks dataset that is larger in size which consists of 1484 images. It gives accuracy of 96.50%.

#### 6.1. Testing model with different images

i. Testing the model by giving unknown input images to model and checking whether it is predicting the image accurately as wearing or not wearing mask.

ii. In order to spot a face in a image, the Haar cascade is employed.

iii. It uses a Haar cascade classifier to read the input image, recognize the face in the image, then crop the only face region in the image.

iv. Resizing the input image size as 224 * 224 and using model.

v. The output is shown as wearing a mask or not wearing mask with the given input image.

### 7. Conclusion

In this project, an effective transfer learning VGG16 is used. The main importance of using vgg16 is that it works well with medium and smaller dataset, less computation and no need of training the model from scratch. After analyzing many review papers, we concluded that pretrained vgg16 gives higher accuracy in smaller dataset which is having 1484 images of 96.50% and less computation than any other Trained Deep Learning algorithms. The same work can be improved further by applying this model to real-time face mask detection and fine tuning the model for medium-sized datasets, which entails adding a custom layer to the model to improve accuracy and implementing a facial recognition system that can be deployed at different workplaces to hold up person detection while wearing the mask.

### CRediT authorship contribution statement

**T.M. Saravanan**: Conceptualization, Methodology, Investigation, Supervision. **K. Karthiha**: Data curation, Writing – original draft. **R. Kavinkumar**: Software, Validation. **S. Gokul**: Visualization, Software, Validation. **Jay Prakash Mishra**: Writing – review & editing.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### References

[1] M. Loey, G. Manogaran, M.H.N. Taha, N.E.M. Khalifa, A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic, Measurement 167 (2021) 108288, https://doi.org/10.1016/j.measurement.2020.108288.

[2] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, J. Hemant, SSDMMV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2, Sust. Cities Soc. 66 (2021) 102692, https://doi.org/10.1016/j.scs.2020.102692.

[3] S. Meivel, K.I. Devi, S.U. Maheswari, J.V. Menaka, Real time data analysis of face mask detection and social distance measurement using Matlab, Mater. Today: Proc., 2021.

[4] N.C. Ristea, R.T. Ionescu, Are you wearing a mask? Improving mask detection from speech using augmentation by cycle-consistent GANs, 2020. arXiv preprint arXiv:2006.10147.

[5] S.V. Militante, N.V. Dionisio, Real-time face mask recognition with alarm system using deep learning, in: 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), IEEE, 2020, pp. 106–110.

[6] A.G. Sandesara, D.D. Joshi, S.D. Joshi, Facial Mask Detection Using Stacked CNN Model, 2020.

[7] M. Jiang, X. Fan, RetinaMask: a face mask detector, 2020. arXiv preprint arXiv:2005.03090.

[8] Y. Chen, M. Hu, C. Hua, G. Zhai, J. Zhang, Q. Li, S.X. Yang, Face mask assistant: Detection of face mask service stage based on mobile phone, IEEE Sens. J. 21 (9) (2021) 11084–11093.

[9] M.S. Ejaz, M.R. Islam, M. Sifatullah, A. Sarker, Implementation of principal component analysis on masked and non-masked face recognition, in: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE, 2019, pp. 1–5.

[10] T. Meenpal, A. Balakrishnan, A. Verma, Facial mask detection using semantic segmentation, in: 2019 4th International Conference on Computing, Communications and Security (ICCCS), IEEE, 2019, pp. 1–5.

[11] Y. Zhang, F. Lv, X. Lu, J. Li, Face detection and alignment method for driver on highway based on improved multi-task cascaded convolutional networks, Multimedia Tools Appl. 78 (18) (2019) 26661–26679.

[12] P. Vallimeena, U. Gopalakrishnan, B.B. Nair, S.N. Rao, CNN algorithms for detection of human face attributes—a survey, in: 2019 International...
Conference on Intelligent Computing and Control Systems (ICCS), IEEE, pp. 576–581.

[14] S.S. Sapkal, A.P. Mengane, T.A. Kalbhor, N.S. Ohol, Human Face Detection and Identification using Deep Metric Learning, 2019.

[15] S. Kumar, S. Singh, J. Kumar, Multiple Face Detection using Hybrid Features with SVM Classifier, in: Data and Communication Networks, Springer, Singapore, 2019, pp. 253–265.

[16] J. Huang, Y. Zhang, H. Chen, Improved Viola-Jones face detection algorithm based on HoloLens, EURASIP J. Image Video Process. 2019 (1) (2019) 1–11.

[17] A. Das, M.W. Ansari, R. Basak, Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV, in: 2020 IEEE 17th India Council International Conference (INDICON). IEEE, 2020, pp. 1–5.

[18] Z. Ning, G. Xutao, Face detection based on skin color extraction scheme, in: IOP Conference Series: Materials Science and Engineering, vol. 569(3), IOP Publishing, 2019, p. 032006.

[19] J.G. Sanders, R. Jenkins, Individual differences in hyper-realistic mask detection, Cogn. Res.: Princ. Impl. 3 (1) (2018) 1–10.

[20] Y. Zhou, D. Liu, T. Huang, Survey of face detection on low-quality images, in: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), IEEE, 2018, pp. 769–773.

Further reading

[12] S. Wu, M. Kan, S. Shan, X. Chen, Hierarchical attention for part-aware face detection, Int. J. Comput. Vis. 127 (6–7) (2019) 560–578.