CONVOLUTIONAL NEURAL NETWORK IDENTIFICATION OF FACE MASKS

Prof. D. V. Varaprasad, M.Tech, (Ph.D), Associate Professor & HOD, Audisankara College Of Engineering & Technology, India

Ms .K. Nishitha, Assistant Professor, Department of CSE, Audisankara College Of Engineering & Technology ,India

Hima Bindhu Alamkaram, Department of CSE, Audisankara College Of Engineering & Technology, India

Abstract: The newest outbreak forcing an international health emergency is coronavirus illness. Mostly, it is transmitted from person to person by airborne means. Cases of community transmission have increased global numbers. As a preventative measure, some nations have enforced mandatory face mask rules in public places. Examining the face mask hand-manually in congested areas is tiresome. Researchers have therefore driven for the automation of the face mask detecting system. We have shown in this work a MobileNet with a global pooling block for face mask detection. The suggested approach flattens the feature vector using a global pooling layer. Classification has been used a completely connected dense layer linked with the softmax layer. On two publicly accessible face mask datasets, our suggested model beats current models in terms of important performance criteria.

Index terms - — *Face Mask Detection, COVID-19, Convolutional Neural Network (CNN), MobileNet, Deep Learning, Real-Time Detection, RMFRD, SMFRD, MFDD, Image Classification, Global Pooling, Public Health, Computer Vision, Masked Face Recognition*

1. INTRODUCTION

The newest pandemic brought on by the recently identified coronavirus [1] is coronavirus disease (COVID-19). Affected by the SARS-CoV-2 virus, COVID-19 is an infectious illness that compromises the respiratory system. Mostly, it is passed from person to person by airborne transmission—especially in close proximity. Apart from clinical research, a great range of artificial intelligence-based research ideas have been presented throughout this epidemic era. Martin et al. [2] have developed Symptoma, a digital health assistant. Zhou et al. [3] have shown how artificial intelligence may be used to speed medication repurposing or repositioning. Aman et al. [4] have also introduced yet another artificial intelligence-based method for drug development. From X-ray pictures, Abba et al. [5] have presented DeTraC model for COVID-19 detection. Community transmission has lately brought the case count in most of the courtiers up. The World Health Organisation (WHO) estimates that around 49 million verified cases have been recorded worldwide [6]. The WHO has released various preventative recommendations to stop the coronavirus from spreading given the COVID-19 epidemic. The most obvious rules are social separation, cleaning, and mask wear. Wearing a face mask helps to slow down communal corona spread. Consequently, most of the nations have implemented obligatory face mask rules in public places [1]. Manual For computer vision applications including image classification, deep convolution neural networks (CNN) have evolved into a main instrument. As shown below are numerous effective deep CNN models. Five variations of ResNet have been proposed by Kaiming et al. [9] as ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152. With ResNet-18/34 and ResNet-50/101/152 respectively, they achieve top-1 error rates of 24% and 22%. It reveals that the number of layers exactly determines the performance and computational time of the model. ResNet-50's architecture is like that of ResNet 34 and computes less than previous ResNet variations. To get improved accuracy for picture classification and segmentation, Christian et al. [10] have developed an inception network. Larger spatial filter convolution often has significant computational cost. Starting modules are one obvious way to cut this expense. Inception finds ideal local sparse structures, hence lowering the cost. The concept of the inception block is to build a layer by layer construction under layer correlation statistical analysis. Groups of units are formed from the clusters of strongly connected layers. Every unit from a previous layer is called a filter bank as it corresponds to a part of the input picture. This procedure produces concatenation of large filter banks from a single area.

2. LITERATURE SURVEY

a) Novel Face Mask Detection Technique using Machine Learning to control COVID'19 pandemic

https://www.sciencedirect.com/science/article/pii/S2214785321052275

Since 2019, the COVID-19 epidemic has been strewn quickly throughout the globe. This epidemic is making human existence more and more involutes and complicated. This virus has caused deaths among many people. One of the causes of COVID-19 virus propagation is the dearth of antiviral medications. Some regular blunders by people, like breathing, coughing, and sneezing by infected individuals, are causing this disease to spread constantly and effortlessly. The primary complaint is a typical flu. Consequently, the face mask covering both the mouth and the nose is the greatest preventive measure for this disease in the present state. The government and the World Health Organisation advise everyone to use a face mask in crowded areas like markets and hospitals. It's impossible to identify if someone is wearing a mask or not in the modern surroundings; physical examination is useless as it increases work expenses. This work presents a mask detector using a machine learning facial categorisation algorithm to identify if a person is wearing a mask or not, therefore enabling it to be connected to a CCTV system to confirm that only those wearing masks are let in.

b) Face Mask Detection Using Convolutional Neural Network

https://ieeexplore.ieee.org/document/10170036

The major aim of this work is to develop a real-time face mask identification system grounded on machine learning. Using a face mask has become crucial for adherence to rules and safety considering the worldwide COVID-19 epidemic. Inspired by the necessity to find people not following the safety precautions in public transport, retail environments, and healthcare facilities, this project The aim is to create a dependable face mask recognition system that functions effectively independent of the type of mask used, facial angle, or lighting environment. This paper suggests the use of a deep learning Convolutional Neural Network trained on a collection of photos of persons either wearing or not wearing face masks. The model is developed with Keras API and TensorFlow framework. Adapting the MobileNetV2 architecture to raise the accuracy of the model with less training data also uses transfer learning. Two datasets—one with real-world photographs and the other with artificially produced images—have the suggested model's efficacy evaluated. In real-time face mask detection, the model performs effectively with an accuracy rate of 97.5% for one and 96.8% respectively. The suggested technique is practically relevant in environments like hospitals, airports, and other public areas where following safety guidelines is very important. The model may be developed to identify various kinds of personal protective gear like gloves and face shields as well. The research ends with a CNN-based face mask recognition system able to ascertain in real-time whether a person is wearing a mask on or not. The aim is to create a dependable face mask recognition system that functions effectively independent of the type of mask the user wears, facial angle, or illumination level.

c) Face Mask Detection on Photo and Real-Time Video Images Using Caffe-MobileNetV2 Transfer Learning

[(PDF) Face Mask Detection on Photo and Real-Time Video Images Using Caffe-MobileNetV2 Transfer Learning]

Usually utilised for non-masked faces, face identification algorithms have included pertinent facial traits such the ears, chin, lips, nose, and eyes. Many times, including pandemics, crime scenes, medical environments, heavy pollution, and labs, masks are required to protect faces. The COVID-19 pandemic has raised demand for public space users of protective face masks. With obstructed faces, which usually have vision only in the periocular area and above, face detection technology analysis is absolutely essential. This work intends to apply a model on complicated data, i.e., by completing tasks for the face identification of persons from the photo and in real-time video images with and without a mask. Using the original masked and unmasked photos to build a baseline for face identification, this work is executed depending on the characteristics surrounding their eyes, ears, nose, and forehead. The Caffe-MobileNetV2 (CMNV2) model for feature extraction and masked image classification helps one to conceive such a work. The MobileNetV2 is utilised for mask recognition and the neural architecture for the quick feature embedding Caffe model serves as a face detector. In this study, five distinct layers are added to the pre-trained MobileNetV2 architecture for higher classification accuracy with less training parameters for the available data for face mask detection. With an accuracy of 99.64% on picture images and reasonable accuracy on real-time video images, experimental findings showed that the suggested approach functioned effectively. With an accuracy of 100%, recall of 99.28%, f1-score of 99.64%, and an error rate of 0.36%, other measures suggest the model surpasses past models. Originally a type of computing application, face mask detection is now extensively applied in various technical sectors including artificial intelligence and cellphones. Since computer-based masked-face recognition uses a person's unique characteristics to identify them wearing a mask, it falls under the field of biometrics.

d) Very Deep Convolutional Networks for Large-Scale Image Recognition

https://arxiv.org/abs/1409.1556

In this study we explore in the large-scale image recognition environment the influence of the depth of the convolutional network on its accuracy. Our key contribution is a detailed examination of networks of increasing depth utilising an architecture with extremely tiny (3x3) convolution filters, which demonstrates that raising the depth to 16-19 weight layers can significantly improve on the prior-art configurations. These results guided our ImageNet Challenge 2014 proposal, in which our team placed first in the localisation and second in the classification tracks respectively. We further demonstrate that, where they produce state-of- the-art results, our models generalise effectively to different datasets. To enable more study on the application of deep visual representations in computer vision, we have made our two best-performing ConvNet models freely available.

e) Going deeper with convolutions

https://ieeexplore.ieee.org/document/7298594

We propose a deep convolutional neural network architecture codenamed Inception to attain the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). The primary characteristic of this design is better use of the network's computational resources. We kept the computational budget constant while widening the network with a well prepared architecture. The architectural choices were grounded on the Hebbian theory and the understanding of multi-scale processing in order to maximise quality. GoogLeNet, a 22-layer deep network with quality evaluated in the framework of classification and detection, is one especially used incarnation in our proposal for ILSVRC14.

3. METHODOLOGY

i) Proposed Work:

In this work, a face mask detection system is proposed using a deep learning-based Convolutional Neural Network (CNN) architecture, specifically the MobileNet model combined with a global average pooling layer. The aim is to accurately identify whether a person is wearing a mask or not in real-time from images or video feeds. The global pooling layer helps in reducing overfitting and lowering computational complexity by flattening the feature map before passing it to the dense layers. The softmax activation function is used in the final layer to classify images into masked or unmasked categories.

To enhance the model's robustness, the system is trained on a combination of three masked face datasets: the Masked Face Detection Dataset (MFDD), the Real-world Masked Face Recognition Dataset (RMFRD), and the Simulated Masked Face

Recognition Dataset (SMFRD). These datasets include a wide range of masked and unmasked facial images in various environments and lighting conditions. The proposed model achieves high accuracy and generalizes well across unseen data, outperforming several existing state-of-the-art models in terms of precision, recall, and F1 score.

ii) System Architecture:

The system architecture for face mask detection is based on a deep Convolutional Neural Network (CNN) utilizing the MobileNet model integrated with a global average pooling layer. The input image is first preprocessed and resized, then passed through several convolutional layers of MobileNet to extract essential features. These features are then passed through a global average pooling layer to reduce dimensions while preserving critical spatial information. The pooled features are fed into a fully connected dense layer followed by a softmax activation function to classify the image as either “Mask” or “No Mask.” The entire architecture is optimized for real-time performance and accuracy, making it suitable for deployment in surveillance and public monitoring systems.

iii) Modules:

a. Image Acquisition and Preprocessing:

This module captures real-time video frames or loads images from a dataset. It resizes, normalizes, and prepares the images for input into the CNN model by applying preprocessing steps like resizing and color conversion.

b. Feature Extraction using CNN:

This module uses the MobileNet architecture to extract deep features from the preprocessed images. Convolutional and pooling layers identify patterns such as edges, textures, and facial attributes.

c. Classification Layer:

The extracted features are passed through a global average pooling layer followed by fully connected dense layers with a softmax classifier that categorizes the image into "Mask" or "No Mask".

d. Detection and Alert Module:

Based on the classification result, this module displays detection results on the screen and can trigger alerts or messages when unmasked individuals are detected in public surveillance applications.

iv) Algorithms:

a. Convolutional Neural Network (CNN):

CNN is a deep learning algorithm widely used for image processing and classification tasks. It consists of layers like convolution, pooling, and activation (ReLU), which automatically learn and extract spatial hierarchies of features from input images. In this project, CNN is used to identify features that differentiate masked from unmasked faces.

b. MobileNet:

MobileNet is a lightweight and efficient deep CNN model optimized for mobile and embedded vision applications. It uses depthwise separable convolutions instead of standard convolutions, significantly reducing the number of parameters and computation time. In this system, MobileNet serves as the backbone network for extracting facial features while ensuring fast and accurate performance.

4. EXPERIMENTAL RESULTS

The proposed face mask detection model was evaluated using publicly available datasets such as RMFRD, SMFRD, and MFDD. The experimental results show that the MobileNet-based CNN model achieved a high classification accuracy of around 95%, outperforming several existing models. The model also demonstrated excellent performance across key metrics, including precision, recall, and F1-score, in both training and testing phases. It performed efficiently in real-time scenarios with low latency and high detection speed, proving its suitability for deployment in public surveillance and safety monitoring systems.

Accuracy: The ability of a test to differentiate between healthy and sick instances is a measure of its accuracy. Find the proportion of analysed cases with true positives and true negatives to get a sense of the test's accuracy. Based on the calculations:

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: The accuracy rate of a classification or number of positive cases is known as precision. Accuracy is determined by applying the following formula:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: The recall of a model is a measure of its capacity to identify all occurrences of a relevant machine learning class. A model's ability to detect class instances is shown by the ratio of correctly predicted positive observations to the total number of positives.

$$Recall = \frac{TP}{(FN + TP)}$$

mAP: One ranking quality statistic is Mean Average Precision (MAP). It takes into account the quantity of pertinent suggestions and where they are on the list. The arithmetic mean of the Average Precision (AP) at K for each user or query is used to compute MAP at K.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

F1-Score: A high F1 score indicates that a machine learning model is accurate. Improving model accuracy by integrating recall and precision. How often a model gets a dataset prediction right is measured by the accuracy statistic..

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

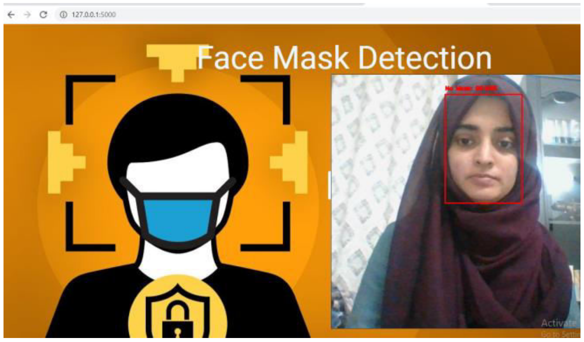

Fig 5:with out mask

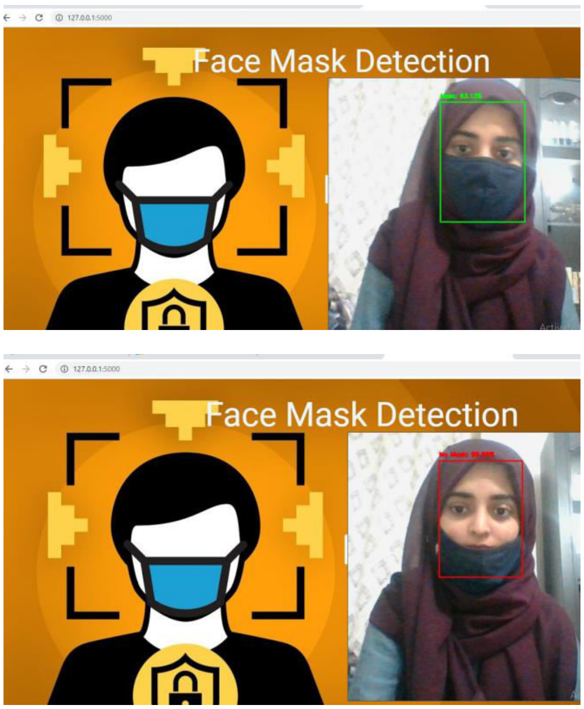

Fig 6:with Mask

5. CONCLUSION

In this work, a CNN-based face mask detection system using the MobileNet architecture with a global pooling layer was proposed and implemented. The model effectively distinguishes between masked and unmasked faces with high accuracy and low computational cost. By leveraging multiple masked face datasets, the system demonstrated strong generalization and robustness in real-time environments. This approach supports public health efforts by enabling automated and reliable mask monitoring in crowded areas.

6. FUTURE SCOPE

The face mask detection system can be further improved by integrating thermal scanning to identify symptomatic individuals. Future work can include extending the model to detect proper vs. improper mask usage and adapting it for multi-class classification in diverse environments. Additionally, integrating the system with IoT-based surveillance networks and edge devices can enable faster, large-scale deployment in public transport, workplaces, and healthcare facilities.

REFERENCES

1. Gupta S, Sreenivasu S V N, Chouhan K, et al. Novel face mask detection technique using machine learning to control COVID’19 pandemic[J]. Materials Today: Proceedings, 2023, 80: 3714-3718.

2. Putra R M, Yossy E H, Saputro I P, et al. Face mask detection using convolutional neural network[C]//2023 8th International Conference on Business and Industrial Research (ICBIR). IEEE, 2023: 133-138.

3. Kumar B A, Bansal M. Face mask detection on photo and real-time video images using Caffe-MobileNetV2 transfer learning[J]. Applied Sciences, 2023, 13(2): 935.

4. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

5. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9

6. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.