International Journal of Innovative Research in Engineering
Volume 3, Issue 2 (March-April 2022), PP: 40-44.
www.theijire.com
ISSN No: 2582-8746

Face Mask Detection

Mohammad Anas[1], Md Sadab Ahmad[2], Rajkishore Yadav[3]
[1,2]Students, Department of Computer Science and Engineering, Institute of Technology and Management, Gorakhpur Uttar Pradesh, India.
[3]Asst. Professor, Department of Computer Science and Engineering, Institute of Technology and Management, Gorakhpur, Uttar Pradesh, India.

How to cite this paper: Mohammad Anas[1], Md Sadab Ahmad[2], Rajkishore Yadav[3], "Face Mask Detection", IJIRE-V3I02-40-44.

Abstract: *The new Coronavirus disease (COVID-19) has seriously affected the world. By the end of November 2021, the global number of new coronavirus cases had already exceeded 29.9 cr and the number of deaths 45,50,000 according to information from the World Health Organization (WHO). To limit the spread of the disease, mandatory face-mask rules are now becoming common in public settings around the world. Additionally, many public service providers require customers to wear face masks in accordance with predefined rules (e.g., covering both mouth and nose) when using public services. These developments inspired research into automatic (computer-vision-based) techniques for face-mask detection that can help monitor public behavior and contribute towards constraining the COVID-19 pandemic. Although existing research in this area resulted in inefficient techniques for face-mask detection, these usually operate under the assumption that modern face detectors provide perfect detection performance (even for masked faces) and that the main goal of the techniques is to detect the presence of face-masks only.*

Key Word: *Covid-19; Detection ; Face Mask; Face Recognition; Neural Network*

I. INTRODUCTION

In December 2019, a new and worryingly contagious primary atypical (viral) pneumo- nia broke out in Wuhan, China. The new disease, called COVID-19, was later found to be caused by a previously unknown zoonotic coronavirus, named SARS-CoV-2. To help limit the spread of this new coronavirus, the World Health Organization (WHO), medical experts as well as governments across the world now recommend that people wear face-masks if they have respiratory symptoms if they are taking care of the people with symptoms or otherwise engage frequently with larger groups of people [1–4]. In response to these developments, research in face mask detection has attracted the attention of the computer vision community recently and initiated efforts towards developing automatic detection Appl. Sci. 2021, 11, 2070 2 of 24models that can help society (through monitoring, screening, and compliance-assessment applications) containing the COVID-19 pandemic.

Face-mask detection represents both detections as well as a classification problem because it requires first the location of the faces of people in digital images and then the decision of whether they are wearing a mask or not. The first part of this problem has been studied extensively in the computer vision literature, due to the broad applicability of face-detection technology. The second part, on the other hand (i.e., predicting whether a face is masked or not), has only gained interest recently, in the context of the COVID-19 pandemic. Although a considerable amount of work has been done over the last year on this part, it typically only tries to detect whether a mask is present in the image. No special attention is given to whether the masks are properly placed on the face and are, hence, worn in accordance with the recommendations of medical experts. This limits the application value of existing face-mask detection techniques and warrants research into computer vision models capable of not only detecting the presence of facial masks in images but also of determining if the masks are worn correctly.

II. MATERIAL AND METHODS

In this study, we revisit these common assumptions and explore the following research questions:
(i) How well do existing face detectors perform with masked-face images? (ii) Is it possible to detect a proper (regulation-compliant) placement of facial masks? (iii) How useful are existing face-mask detection techniques for monitoring applications during the COVID-19 pandemic?

To answer these and related questions we conduct a comprehensive experimental evaluation of several recent face detectors for their performance with masked-face images. Furthermore, we investigate the usefulness of multiple off the-shelf deep-learning models for recognizing correct face-mask placement. Finally, we design a complete pipeline for recognizing whether face masks are worn correctly or not and compare the performance of the pipeline with standard face mask detection models from the literature. To facilitate the study, we compile a large dataset of facial images from the publicly available MAFA and

Wider Face datasets and annotate it with compliant and non-compliant labels. The annotation dataset, called Face-Mask-Label Dataset (FMLD), is made publicly available to the research community.

Study Design: Prospective open label observational study
Study Location: This was a face mask detection based study done in Department of Computer Science , at Institute of Technology and Management, Gorakhpur, Uttar Pradesh.

Inclusion criteria:

1. Diabetic patients (fasting blood glucose ≥ 126 mg/dL [7.0mmol/L])
2. Either sex
3. Aged ≥ 18 years,
4. Patients have a total cholesterol level of ≥154.68 mg/dl , LDL-C 96.6 mg/dl, HDL-C ≤ 138.6 in men and ≤46.3 mg/dl in women.
5. Fasting triglycerides ≥ 150.56 mg/dl, obtained within 1 week before the first use of statins which was then compared at first- and second-year intervals.

Procedure methodology
There are two main approaches for Face Detection:

Feature Base Approach
Image Base Approach

Feature Base Approach
Objects are usually recognized by their unique features. There are many features in a human face, which can be recognized between a face and many other objects. It locates faces by extracting structural features like eyes, nose, mouth etc. and then uses them to detect a face. Typically, some sort of statistical classifier qualified then helpful to separate between facial and non-facial regions. In addition, human faces have particular textures which can be used to differentiate between a face and other objects. Moreover, the edge of features can help to detect the objects from the face. In the coming section, we will implement a feature-based approach by using Open CV.

Image Base Approach
In general, Image-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images. The learned characteristics are in the form of distribution models or discriminant functions that is consequently used for face detection. In this method, we use different algorithms such as Neural-networks, HMM, SVM, AdaBoost learning. In the coming section, we will see how we can detect faces with MTCNN or Multi-Task Cascaded Convolutional Neural Network, which is an Image-based approach of face detection

Face detection algorithm

One of the popular algorithms that use a feature-based approach is the Viola-Jones algorithm and here I am briefly going to discuss it. If you want to know about it in detail, I would suggest going through this article, Face Detection using Viola Jones Algorithm.
Viola-Jones algorithm is named after two computer vision researchers who proposed the method in 2001, Paul Viola and Michael Jones in their paper, “Rapid Object Detection using a Boosted Cascade of Simple Features”. Despite being an outdated framework, Viola-Jones is quite powerful, and its application has proven to be exceptionally notable in real-time face detection. This algorithm is painfully slow to train but can detect faces in real-time with impressive speed.
Given an image(this algorithm works on grayscale image), the algorithm looks at many smaller subregions and tries to find a face by looking for specific features in each subregion. It needs to check many different positions and scales because an image can contain many faces of various sizes. Viola and Jones used Haar-like features to detect faces in this algorithm.
Face detection and Face Recognition are often used interchangeably but these are quite different. In fact, Face detection is just part of Face Recognition.
Face recognition is a method of identifying or verifying the identity of an individual using their face. There are various algorithms that can do face recognition but their accuracy might vary. Here I am going to describe how we do face recognition using deep learning.
As mentioned before, here we are going to see how we can detect faces by using an Image-based approach. MTCNN or Multi-Task Cascaded Convolutional Neural Network is unquestionably one of the most popular and most accurate face detection tools that work this principle. As such, it is based on a deep learning architecture, it specifically consists of 3 neural networks (P-Net, R-Net, and O-Net) connected in a cascade.
So, let’s see how we can use this algorithm in Python to detect faces in real-time. First, you need to install MTCNN library which

contains a trained model that can detect faces.

Face Recognition

Face detection and Face Recognition are often used interchangeably but these are quite different. In fact, Face detection is just part of Face Recognition.

Face recognition is a method of identifying or verifying the identity of an individual using their face. There are various algorithms that can do face recognition but their accuracy might vary. Here I am going to describe how we do face recognition using deep learning.

In fact here is an article, Face Recognition Python which shows how to implement Face Recognition.

Statistical analysis

Like object detection, face detection adopts the same architectures as one-stage and two-stage detectors, but in order to improve face detection accuracy, more face-like features are being added. However, there is occasional research focusing on face mask detection. Some already existing face mask detectors have been modeled using OpenCV, Pytorch Lightning, MobileNet, RetinaNet and Support Vector Machines. Here, we will be discussing two projects. One project used Real World Masked Face Dataset (RMFD) which contains 5,000 masked faces of 525 people and 90,000 normal faces [8]. These images are 250 x 250 in dimensions and cover all races and ethnicities and are unbalanced. This project took 100 x 100 images as input, and therefore, transformed each sample image when querying it, by resizing it to 100x100. Moreover, this project uses PyTorch then they convert images to Tensors, which is the base data type that PyTorch can work with. RMFD is imbalanced (5,000 masked faces vs 90,000 non-masked faces). Therefore, the ratio of the samples in train/validation while splitting the dataset was kept equal using the train test split function of sklearn. Moreover, to deal with unbalanced data, they passed this information to the loss function to avoid unproportioned step sizes of the optimizer. They did this by assigning a weight to each class, according to its representability in the dataset. They assigned more weight to classes with a small number of samples so that the network will be penalized more if it makes mistakes predicting the label of these classes. While classes with large numbers of samples, they assigned to them a smaller weight. This makes their network training agnostic to the proportion of classes. The weights for each class were chosen using the formula below: Class W eight = 1 − Class Cardinality Cardinalities of all classes To load the data efficiently this project used the data loader. For instance, in this project, they used the PyTorch lighting, and to load them for training and validation they divided data into 32 batches and assigned the works of loading to the 4 number of workers, and this procedure allowed them to perform multi-process data loading. Like most of the projects, this project also used Adam optimizer. If any Model has a high rate of learning, it learns faster, but it bounces a lot to reach the global minima and may diverge from the global minima. However, a small learning rate may take considerably lower time to train, but it reaches to the global minima. If the loss of the model declines quickly for any learning rate, then that learning rate would be the best learning rate. However, it seems that this project considered the 0.00001 learning rate would be the best for their model so that it could work efficiently. To train the model they defined a model checkpointing callback where they wanted to save the best accuracy and the lowest loss. They tried to train the model for 10 epochs and after finding optimal epoch, they saved the model for 8 epochs to test on the real data. To get rid of the problem of occlusions of the face which causes trouble face detectors to detect masks in the images, they used a built-in OpenCV deep learning face detection model. For instance, the Haar-Cascade model could be used but the problem of the Haar-Cascade model is that the detection frame is a rectangle, not a square. That is why, without capturing the portion of the background, the face frame can fit the entirety of the face, which can interfere with the face mask model predictions. In the second project [9], a dataset was created by Prajna Bhandary using a PyImageSearch reader. This dataset consists of 1,376 images belonging to all races and is balanced. There are 690 images with masks and 686 without masks. Firstly, it took normal images of faces and then created a customized computer vision Python script to add face masks to them. Thereby, it created a realworld applicable artificial dataset. This method used the facial landmarks which allow them to detect the different parts of the faces such as eyes, eyebrows, nose, mouth, jawline etc. To use the facial landmarks, it takes a picture of a person who is not wearing a mask, and, then, it detects the portion of that person's face. After knowing the location of the face in the image, it extracted the face Region of Interest (ROI). After localizing facial landmarks, a picture of a mask is placed into 3 the face. In this project, embedded devices are used for deployment that could reduce the cost of manufacturing. MobileNetV2 architecture is used as it is a highly efficient architecture to apply on embedded devices with limited computational capacity such as Google Coral, NVIDIA Jetson Nano. This project performed well, however, if a large portion of the face is occluded by the mask, this model could not detect whether a person is wearing a mask or not. The dataset used to train the face detector did not have images of people wearing face masks as a result, if the large portion of faces is occluded, the face detector would probably fail to detect properly. To get rid of this problem, they should gather actual images of people wearing masks rather than artificially generated images.

III. RESULT

This proposed work utilizes a consecutive Convolutional Brain Network for distinguishing and perceiving human appearances of people with cover or without it. CNN model what's more, Haar Cascade Algorithm works with programmed identification and acknowledgment of human face which survive the commotion varieties and foundation varieties brought about by the encompassing and give more exact and exact result. It likewise helps toovercome the lopsided idea of the latest thing of face acknowledgment and identification. From the tests plainly the proposed CNN accomplishes a high precision when contrasted with different designs. The proposed calculation turns out successfully for various sorts of pictures. These outcomes recommend that the proposed CNN model lessens intricacy and make technique computationally compelling. The proposed framework works well really for grayscale as well concerning the shading picture with veils on it or without covers on it.

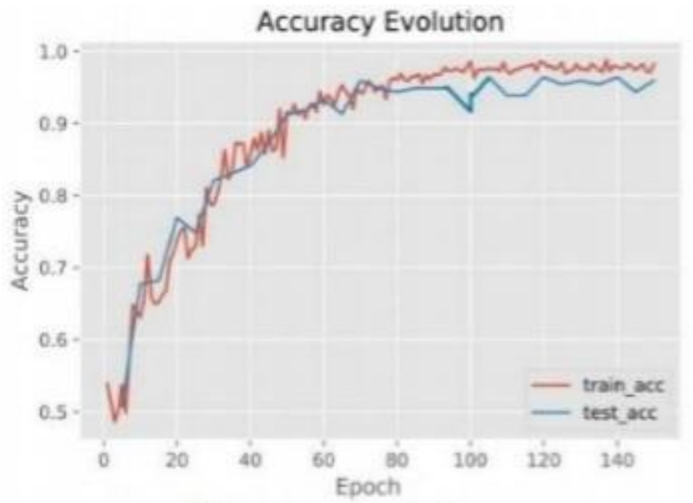

Fig3. Accuracy analysis

	Precision	Recall	F1-Score	Support
With mask	0.99	0.99	0.99	138
Without Mask	0.99	0.99	0.99	138
Accuracy			0.99	276
Macro Avg	0.99	0.99	0.99	276
Weighted Avg	0.99	0.99	0.99	276

Table1. Performance Analysis

IV. DISCUSSION

Over the period there have been many advancements in the deep learning towards object detection and recognition in various application domains (27,18]. In general, most of the works focus on image reconstruction and face recognition for identity verification. But the main aim of this work is to identify people who are not wearing masks in public places to control the further transmission of COVID-19. Bosheng Qin and Dongxiao Li (22) have designed a face mask identification method using the SRCNet classification network and achieved an accuracy of 98.7% in classifying the images into three categories namely "correct facemask wearing", "incorrect facemask wearing" and "no facemask wearing". Md. Sabbir Ejaz et al. [7] implemented the Principal Component Analysis (PCA) algorithm for masked and un-masked facial recognition. It was noticed that PCA is efficient in recognizing faces without a mask with an accuracy of 96.25% but its accuracy is decreased to 68.75% in identifying faces with a mask. In a similar facial recognition application, Park et al. (12) proposed a method for the removal of sunglasses from the human frontal facial image and reconstruction of the removed region using recursive error compensation.

Li et al. [14] used YOLOv3 for face detection, which is based on deep learning network architecture named darknet-19, where WIDER FACE and Celebi databases were used for training, and later the evaluation was done using the FDDB database. This model achieved an accuracy of 93.9%. In a similar research, Nizam et al. [6] proposed a GAN based network architecture for the removal of the face mask and the reconstruction of the region covered by the

mask. Rodriguez et al. [17] proposed a system for the automatic detection of the presence or absence of the mandatory surgical mask in operating rooms. The objective of this system is to trigger alarms when a staff is not wearing a mask. This system achieved an accuracy of 95%. Javed et al. [13] developed an interactive model named MRGAN that removes objects like microphones in the facial images and reconstructs the removed region's using a generative adversarial network. Hussain and Balushi [11] used VGG16 architecture for the recognition and classification of facial emotions. Their VGG16 model is trained on the KDEF database and achieved an accuracy of 88%. Following from the above context it is evident that specially for mask detection very limited number of research articles have been reported till date whereas further improvement is desired on existing methods. Therefore, to contribute in the further improvements of face mask recognition in combat against COVID-19.

V. CONCLUSION

Over the period there have been many advancements in the deep learning towards object detection and recognition in various application domains (27,18]. In general, most of the works focus on image reconstruction and face recognition for identity verification. But the main aim of this work is to identify people who are not wearing masks in public places to control the further transmission of COVID-19. Bosheng Qin and Dongxiao Li (22) have designed a face mask identification method using the SRCNet classification network and achieved an accuracy of 98.7% in classifying the images into three categories namely "correct facemask wearing", "incorrect facemask wearing" and "no facemask wearing". Md. Sabbir Ejaz et al. [7] implemented the Principal Component Analysis (PCA) algorithm for masked and un-masked facial recognition. It was noticed that PCA is efficient in recognizing faces without a mask with an accuracy of 96.25% but its accuracy is decreased to 68.75% in identifying faces with a mask. In a similar facial recognition application, Park et al. (12) proposed a method for the removal of sunglasses from the human frontal facial image and reconstruction of the removed region using recursive error compensation.

Li et al. [14] used YOLOv3 for face detection, which is based on deep learning network architecture named darknet-19, where WIDER FACE and Celebi databases were used for training, and later the evaluation was done using the FDDB database. This model achieved an accuracy of 93.9%. In a similar research, Nizam et al. [6] proposed a GAN based network architecture for the removal of the face mask and the reconstruction of the region covered by the
mask. Rodriguez et al. [17] proposed a system for the automatic detection of the presence or absence of the mandatory surgical mask in operating rooms. The objective of this system is to trigger alarms when a staff is not wearing a mask. This system achieved an accuracy of 95%. Javed et al. [13] developed an interactive model named MRGAN that removes objects like microphones in the facial images and reconstructs the removed region's using a generative adversarial network. Hussain and Balushi [11] used VGG16 architecture for the recognition and classification of facial emotions. Their VGG16 model is trained on the KDEF database and achieved an accuracy of 88%. Following from the above context it is evident that specially for mask detection very limited number of research articles have been reported till date whereas further improvement is desired on existing methods. Therefore, to contribute in the further improvements of face mask recognition in combat against COVID-19.

References

1. Dataset.https://github.com/prajnasb/observations.Accessed 25 May 2020
2. Paris tests face-mask recognition software on metro riders. http://bloomberg.com/.Accessed 25 May 2020
3. Who coronavirus disease (COVID-19) dashboard. https://covid19.who.int/. Accessed 25 May 2020
4.Agarwal, S., Punn, N.S., Sonbhadra, S.K., Nagabhushan, P., Pandian, K., Saxena, P.: Unleashing the power of disruptive and emerging technologies amid COVID 2019: a detailed review. arXiv preprint arXiv:2005.11507 (2020)
5.Chollet, F.: Xception: deep learning with depthwise separable convolutions. CORR abs/1610.02357 (2016). http://arxiv.org/abs/1610.02357 6. 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)