Face mask recognition during KYC generation from a live photo detection methodology

Debalina Barua
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
debalina.barua@g.bracu.ac.bd

Mumtahina Ahsan
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
mumtahina.ahsan@g.bracu.ac.bd

Moumita Khandaker
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
moumita.khandaker@g.bracu.ac.bd

H M Zahidul Amin
Department of Computer Science and Engineering
BRAC University
66 Mohakhali, Dhaka 1212, Bangladesh
muhammed.zahidul.amin@g.bracu.ac.bd

***Abstract*—Post Pandemic world of Covid-19 has set human race to a transitioned frequency. Through this transitional period the world needs to serve itself with much needed technology and services. The need of time is now to build the system compatible to the crisis we face together that ensures a risk-free safe environment. New regulations and measures have been established in order to provide safety that includes regular wear of face mask. It is necessary to strictly execute this new rule in public places that helps reducing the spread of virus. Along the mask, a mask detector system comes hand in hand playing a vital role. On official image recording sectors the mask detection is a requirement to ensure the identity of a personnel. This surveillance method enables the alert system to remove the mask while taking any photograph for documentation that demands clear photographic identity. In this paper we have briefly discussed this approach of face mask detection system while capturing a photo using machine learning algorithms.**

***Index Terms*—machine learning, photographic identity, face mask detection**

I. Introduction

Since late 2019 to till now, Covid-19 pandemic has its effects wide spread in all over the world. Sourcing from China's Wuhan it traveled its way to 114 countries of the world recognized as a fatal disease by World Health Organization. With tireless work and dedication of the medical professionals, researchers, scientists, biologists the world got the vaccine introduced by 2 years. This virus travels on air and spreads very easily through an infected person's nearby physical contact. Alike social distancing and hand sanitization, face mask has been an essential tool to restrict the spread of this virus to certain extent. But some primary issues also occurred along with these safety measures. In the field of image processing, to identify a face mask from an image has been equally crucial since then. Face detection has various use cases ranging from face recognition to capturing facial motions, where the latter calls for the face to be revealed with very high precision. Due to the rapid advancement in the domain of machine learning algorithms, the jeopardies of face mask detection technology seem to be well addressed yet. This technology is more relevant today because it is used to detect faces not only in static images and videos but also in real-time inspection and supervision. We have applied TFLite Model to the dataset of images containing facial structure with or without face mask worn. The execution of this procedure detects mask or no mask images that helps the documentation or recording of images captured for official works that requires a clear face image without a mask.

II. Literature Review

If we shed light to the previously done work discussion, several datasets on facial structure detection can be found that are used in building face mask identifying models. The difference between previous datasets and current are the photos collected were taken from supervised surroundings and now they are mostly taken live using systems like MALF, Wider Face , Celeb A , IJB-A etc [1], [2], [3]. There can be many variety of analyzing models to detect a mask on face.

In [4], there were boosted cascades including easy haar features and usage of Viola-Jones face detector which showed a classification based on boosting. Farther researches have the implementation of multiview face mask detector that is evolved from the Viola-Jones model in [5]. Later decision tree algorithm was also introduced for detection. [5] did some similar detection activity using a DPM detector with 97.14% accuracy that was remarkable. The Random forest classification tree model applied in Ramanan's 2006 research work could guess the facial structure and poses precisely. Although these DPM models show great precision in result the calculation process is very costly. These are classified on the basis of Deformable Part model and implemented at around 30 thousand image dataset. The bigger the dataset the easier and more accurate the detection process is and it gives the training and testing data a large scale of flexibility. [6] shows some detection analysis of face mask as well. We can see algorithms learning from data provided by users directly on classifications based on CNN [7] that also applies deep learning. [8] similarly applied cascade CNN and [9]. A CNN structure on Contextual Multi-Scale Region was introduced on [10] that left an implacable impression. [11] on the contrary has reduced the amount of error in the previously mentioned model by developing biased obstruction on the grid loss layer. [12] brought up the 3D version of that CNN model that took in observation of facial structure.

III. Methodology

A. Dataset Description

The dataset that we have used for our project is a public dataset named Face Mask Lite Dataset that has been taken from the platform Kaggle. The dataset contains 10,000 HD images of people wearing masks as well as without wearing masks. All the images used in this dataset were generated using style GAN-2.

Fig. 1. Face Mask Lite Dataset

B. TensorFlow Lite

TensorFlow Lite is a framework of software packages, which enables on-device Machine Learning. It is an open-source deep learning framework. TensorFlow Lite framework allows developers to run their models on targeted hardware such as: mobile systems (like- Android or iOS), Linux based embedded systems (like- Raspberry Pi or Microcontrollers), and edge devices. As TensorFlow Lite provides high performance, it is very useful for model optimization. Some very popular and common machine learning tasks such as image classification, text classification, object detection, pose estimation, question answering, etc can be done by using this framework. In this paper, we proposed face mask recognition using live photo detection methodology. To develop our proposed model we have generated it using the TensorFlow Lite model.

C. Generating TensorFlow Lite model

For generating the TensorFlow Lite model at first we chose a machine learning model. As we have worked on face mask recognition, we have chosen Binary Classification Model. Then we trained our model. After training the model, we converted the model to the Tensorflow Lite model in the .tflite format. The .tflite format is known as a Flat Buffer file, which is common to the TensorFlow Lite framework and compatible hardware. Flat Buffer model provides several advantages over TensorFlow's protocol, like as- reducing size. And lastly, we deployed the Tensorflow Lite Flat Buffer file on a mobile device running on iOS and Android.

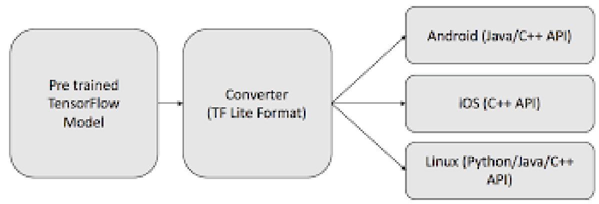

Fig. 2. Working Process of Tensor Flow

D. MobileNet

One of the smallest Deep Neural Networks that can operate on hardware without powerful GPUs is MobileNet. It is quick and effective. When utilizing a framework like Keras, the implementation of neural networks is quite straightforward (on TensorFlow). MobileNet is a streamlined architecture that uses depth wise separable convolutions to construct lightweight deep convolutional neural networks and provides an efficient model for mobile and embedded vision applications [13]. Rather than using the standard convolution method used by conventional CNN(Convolutional neural networks) , it makes use of the concepts of depth and point convolution. This makes CNN more effective at predicting images, allowing them to compete in mobile systems as well. We are utilizing them as our image recognition model because they greatly cut the comparison and recognition times while yet delivering better results quickly.

IV. Implementation

There are five steps involved in the detection of face masks and also not detection of face masks as shown in Fig 2.

With the aid of machine learning, we hope to identify the provided images in our dataset for binary classification. In order to recognize photos, we're assuming that we already have a pre-trained model in Tensorflow. In order to better detect images and anticipate them using coordinates and indexing, we will import architectures using Keras and Tensorflow. We used the tflite model format as we are going to use it on mobile devices. Because with tensorflow lite framework we can embed the model in an App project. Google's TensorFlow—that ease the process of acquiring data, training models, serving predictions, and refining future results. Here we use the tensorflow lite framework and passed the kaggle dataset image to the model to detect the class and model will give the class. TensorFlow is an end-to-end open source platform for machine learning which is a flexible ecosystem of tools, libraries, and community resources to build and deploy ML-powered

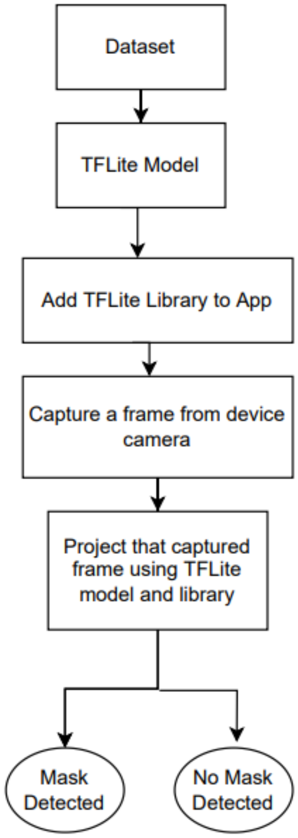

Fig. 3. Face Mask Lite Dataset

applications. We used TensorFlow Lite which is a mobile library for deploying models on mobile, microcontrollers and other edge devices. This tflite also supports multiple platforms like android and iOS. Then we have passed the image to the model to detect the class. Then model will give the class probability by which we can select a threshold of class detection. In this case we set the threshold as 60 % MobileNet used 28 layers Convolutional Neural Networks for Mobile Vision Applications as shown in Fig 3. A convolutional Neural Network (CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

V. Result Analysis

Our application can clearly detect in real time when the person is wearing a mask or not when doing Know Your Customer (KYC). The model is tested using a set of steps to ensure that it is correctly predicting data. On the basis of the testing data set, predictions are made as the first step. While training the model, loss and accuracy are recorded after

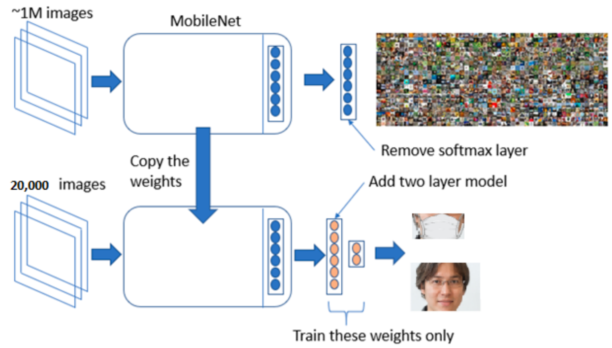

Fig. 4. Schematic diagram of the MobileNet model

specific iterations. Accuracy keeps rising while loss keeps falling, according to the model's training outcomes. There comes a point where additional iterations are not necessary since the accuracy is steady.

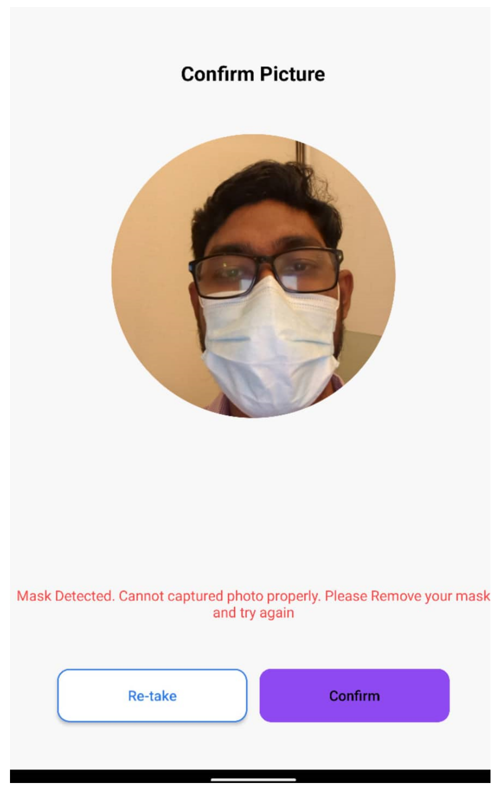

Fig. 5. Mobile app detecting face with mask on

VI. Conclusion

This research presents an effective CNN model for Real-time Facemask Detection based on MobileNet. The suggested method can assess whether or not a mask is being worn properly in real-time video streams, and it demonstrated accuracy in training and testing. Numerous tests are run to demonstrate the MobileNet model's successful detection of facemasks in real-time movies. As future work, further experiments will be conducted to evaluate the performances of the proposed solution. In addition, we plan to implement the proposed solution in real world surveillance cameras in public areas to check if people are following rules and wearing masks appropriately. We also want to train our model so that it can detect images of people in different lighting conditions to get better results. In the dataset that we have used to train our model, the images contained people wearing the same type of masks. Hence, we would like to try to detect masks of different colours as well.

REFERENCES

[1] Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Wider face: A face detection benchmark. 11 2015.

[2] Brendan Klare, Benjamin Klein, Emma Taborsky, Austin Blanton, Jordan Cheney, Kristen C. Allen, Patrick Grother, Alan Mah, Mark Burge, and Anil K. Jain. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1931–1939, 2015.

[3] Binh Yang, Junjie Yan, Zhen Lei, and S. Li. Fine-grained evaluation on face detection in the wild. *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 1:1–7, 2015.

[4] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. volume 1, pages I–511, 02 2001.

[5] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, 2016.

[6] Dong Chen, Shaoqing Ren, Yichen Wei, Xudong Cao, and Jian Sun. Joint cascade face detection and alignment. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 109–122, Cham, 2014. Springer International Publishing.

[7] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 06 2015.

[8] Haoxiang Li, Zhe Lin, Xiaohui Shen, and Jonathan Brandt. A convolutional neural network cascade for face detection. pages 5325–5334, 06 2015.

[9] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[10] Chenchen Zhu, Yutong Zheng, Khoa Luu, and Marios Savvides. *CMS-RCNN: Contextual multi-scale region-based CNN for unconstrained face detection*, pages 57–79. 08 2017.

[11] Michael Opitz, Georg Waltner, Georg Poier, Horst Possegger, and Horst Bischof. Grid loss: Detecting occluded faces. 09 2016.

[12] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. A convolutional neural network cascade for face detection. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5325–5334, 2015.

[13] Andrew Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. 04 2017.