Real Time Face Mask Detection

Saumya Srivastava
Artificial Intelligence & Machine Learning Engineering,
RV College of Engineering® ,
Bengaluru, India

Sabaa Namazi
Artificial Intelligence & Machine Learning Engineering
RV College of Engineering®
Bengaluru, India

Gayathri BR
Artificial Intelligence & Machine Learning Engineering
RV College of Engineering®
Bengaluru, India

Abstract— **Real time face mask detection model is designed to find whether the people are wearing masks or not . This model uses Depp Learning as well as computer vision. Deep Learning part includes Convolutional Neural Network which is MobileNetv2 and Computer Vision includes Open CV. After the covid pandemic it has become kind of necessary to wear masks in public places. To make surveillance easy it is need to have such kind of models in real world. Through CCTVs we should be able to detect who all are wearing masks. In this model MobileNetV2 is a light weight CNN model used for mask classification whereas OpenCV is used for face detection. The dataset has images labeled as with masks and without masks. Tensor Flow , Keras and OpenCV work together to make this model .The training process involves data augmentation , binary classification and optimization. Pre-Processing, Data augmentation, training , testing, batch normalization, one-hot encoding are some of the processes which is done in this project. Adam is used as an optimizer in this architecture . The dataset has 1915 images of people wearing mask and 1918 of people not wearing mask. Since this is a binary classification model, it will show green box if it can correctly predict that the person is wearing mask and red if the person is not wearing mask.This model is a boon for surveillance.**

Keywords—Face Mask Detection, Real time, CNN (Convolutional Neural Network), DNN (Deep Neural Networks), Computer Vision(CV),Data augmentation, Data set

I. INTRODUCTION

After the rise of coronavirus it has become kind of mandatory to wear mask in public places to stop the spread of virus. Real time monitoring of large amount of people to distinguish between the people wearing masks and not wearing masks is a difficult job. Manual monitoring is not possible for huge crowd. After the pandemic wearing a mask was kind of compultion to be followed.Traditional approach is not efficient and is prone to error . These are the reasons for the need of deployment of a deep learning real time face detection model using a light Convolutional Neural Network which is MobileNetV2 along with it integrates pre trained face detection model (SSD ResNet) from OpenCV which is used for face localization. MobileNetV2 then classifies faces as wearing mask or not wearing mask. By deploying a model which has OpenCv for video streaming and Tensor Flow/ Keras for deep learning it is efficient , scalable and reliable method to be used in public spaces for monitoring, workplace surveillance. Public spaces like hospitals, airports , traffic signals, schools, colleges and offices where wearing masks is necessary , this model holds great importance . This project main aim is to ensure public health is maintained by automating the task and reducing human involvement. Even a single person can cause the spread of these type of diseases therefore, it is important to have a system that ensures that virus is not spread , one way of this is to make sure everyone is wearing masks.

II. LITERATURE SURVEY

C. Jagadeeswar[1] presents Performance Evaluation of Intelligent Face Mask Detection System with various Deep Learning Classifiers The system is evaluated with different classifiers. Different classifiers like MobileNetV2, RESNET50, VGG16, each of them was compared with Optimizers like ADAM, ADAGRAD, SGD to yield the highest accuracy. ADAM gave maximum accuracy. The best framework might be executed alongside interfacing with caution and alarming frameworks soon. This framework might be incorporated with a framework which can coordinate with a framework actualizing social distancing which can make it a healthy framework that can welcome emotional effect on the spread. Vinitha.V1, Velantina.V2 [2] presents Covid-19 facemask detection with deep learning and computer vision. The system comprises of MobileNet as the spine which can be very well utilized for high and low calculation situations. In order to extract more robust features, learning is used to gain weights from a similar task face detection, which is trained on large datasets. The proposed method accomplishes state-of-art results on public face mask dataset. Wei Bu, Jiangjian Xiao, Chuanhong Zhou, Minmin Yang, Chengbin Peng [3] presents A Cascade Framework For Masked Face Detection. This algorithm is based on a recently planned CNN course structure comprises of three CNNs. Additionally, another dataset called “MASKED FACE dataset” is proposed which have 160 images for training and 40 images for testing. To defeat the overfitting issue due to lack of training samples, the model is pre-trained with WIDER FACE dataset, and finetuned them with MASKED FACE training set. Overfitting cab be prevented if the epochs or training is stopped once it is stabilized The system can be used in CCTV footages to identify whether a person is wearing a mask correctly so that he does not pose any hazard to others Amit Chavda, Jason Dsouza, Sumeet Badgujar, Ankit Damani [4] presents Multi-Stage CNN Architecture for Face Mask Detection. A two-stage face mask detector was introduced. First stage uses pre-trained RetinaFace model for face detection, after comparing its performance with Dlib and MTCNN. Second stage uses NASNetMobile based model for classifying faces as “masked” or “unmasked”. Besides, centroid tracking is used to improve performance on video streams. At present, the model gives 5 FPS inference speed

979-8-3503-1502-8/23/$31.00 ©2023 IEEE

on a CPU. We plan to improve this up to 15 FPS, making our model deployable for CCTV cameras, without need of a GPU. Stage 1 and stage 2 models can be easily replaced with improved models that would give better accuracy and lower latency Aniruddha Srinivas Joshi, Shreyas Srinivas Joshi, Goutham Kanahasabai, Rudraksh Kapil, Savyasachi Gupta," Deep Learning Framework to Detect Face Masks from Video Footage [5] presents Deep Learning Framework to Detect Face Masks from Video Footage. A profoundly successful face identification model is utilized for getting facial pictures and signals .The subsequent methodology is very strong and is assessed on a custom dataset that got for this work. The proposed approach was discovered to be successful as it depicted high exactness, review, and precision esteems on the picked dataset which contained videos with fluctuating occlusions and facial angles. Mass screening is little difficult in crowded places like railway stations, bus stops, streets, schools, colleges, etc. so the system yielding better accuracy can be created. Reny Jose [6] presents A convolutional neural network (CNN) approach to detect face using tensorflow and keras. A deep convolutional neural network (CNN) is utilized to extricate highlights from input pictures. Keras is utilized for actualizing CNN additionally Dlib and OpenCV for adjusting faces on information pictures. Face acknowledgment execution is assessed utilizing a custom dataset. Facial acknowledgment is a powerful apparatus that can help law makers perceive lawbreakers and software organizations are utilizing the innovation to help clients access their innovation. This innovation can be additionally evolved to be utilized in different avenues like ATMs, getting to private records, or other sensitive materials. This can make other safety efforts, for example, passwords and keys obsolete. Another way that innovators are looking to execute facial acknowledgment is inside subways and other transportation outlets. They are hoping to use this innovation to use faces as credit cards to pay for our transportation charge. Adithya K1, Jismi Babu2 [7] presents A Review on Face Mask Detection using Convolutional Neural Network. By the improvement of AI and image processing analysis present strategies for mask detection. This paper gives tells about artificial intelligence role in detecting the issue. By utilizing image processing analysis and AI technique is utilized for finding out mask detection. Face mask identification can be done through different strategies. Essentially convolutional neural network technique is utilized quickly. The precision and decision making are exceptionally high in CNN contrasted with others. An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network [8] presents An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network. In this work, a framework is proposed that confine the development of Coronavirus by discovering individuals who are not wearing any facial mask in a smart city network where all the public spots are checked with Closed-Circuit Television (CCTV) cameras. While an individual without a mask is identified, the corresponding authority is informed through the city network. A deep learning design is prepared on a dataset that comprises of pictures of individuals with and without masks gathered from different sources. The created framework faces challenges in arranging faces covered by hands since it nearly resembles the individual wearing a mask. While any individual without a face mask is going on any vehicle, the framework can't find that individual accurately. For a thickly populated zone, recognizing the face of every individual is troublesome. For this sort of situation, recognizing individuals without face mask would be very hard for the proposed framework. To get the best result out of this framework, the city should have an enormous number of CCTV cameras to screen the entire city as well as dedicated manpower to uphold appropriate laws on the violators. Since the data about the violator is sent through SMS, the framework fails when there is an issue in the network. Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, Heling Chen, Yu Miao, Zhibing Huang, and Jinbi Liang [9] presents Masked Face Recognition Dataset and Application. It is urgent to improve the acknowledgment execution of the current face recognition innovation on the masked faces. Most current progressed face acknowledgment approaches are planned dependent on deep learning, which rely upon an enormous number of face samples. Notwithstanding, at present, there are no freely accessible masked face recognition datasets. To this end, this work proposes three sorts of masked face datasets, including Masked Face Detection Dataset (MFDD), Real world Masked Face Recognition Dataset (RMFRD) and Simulated Masked Face Recognition Dataset (SMFRD). Among them, to the best of our knowledge, RMFRD is as of now the world's biggest real-world masked face dataset. These datasets are freely accessible to industry and the academia, based on which different applications on masked faces can be created. The multigranularity masked face recognition model created accomplishes almost 95% accuracy, exceeding the outcomes reported by the industry. MFDD, RMFRD and SMFRD datasets are built, and built up a state-of-the-art algorithm dependent on these datasets. The calculation will serve the utilizations of contactless face authentication in local area access, campus management, and enterprise resumption scenarios. This exploration has contributed logical and innovative capacity to the avoidance and control of Covid epidemics and the resumption of creation in industry. Moreover, because of the continuous event of haze weather, individuals will frequently wear covers, and the requirement for face recognition with mask will continue for quite a while. Mohamed Loey, Gunasekaran Manogaran b,c , Mohamed Hamed N. Taha d ,Nour Eldeen M. Khalifa [10] presents Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. The target of this paper is to comment on and confine the clinical face mask objects, all things considered, pictures. Wearing a clinical face mask in open territories, ensure individuals from COVID-19 transmission among them. The accomplished results presumed that the adam optimizer accomplished the most elevated normal accuracy level of 81% as a detector. As a further study, it is intended to distinguish a kind of masked faces in picture and video based on deep learning models.

III. DESIGN AND IMPLEMENTATION

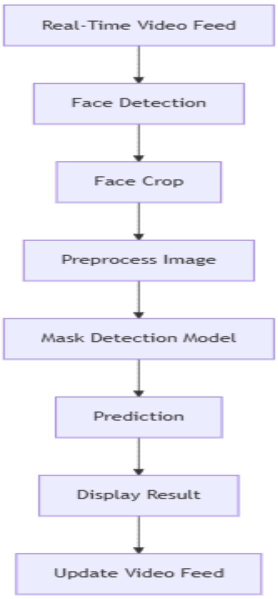

Fig.1. Block diagram of Face Mask Detection

Figure 1 shows the block diagram of Face Mask Detection in real time . The first step is to capture video in real time through web cam . OpenCV model ResNet is used to locate the faces in video stream. The detected faces are then cropped.Then image processing takes place in which the image is resized(224*224) pixels , normalized and converted to a format which is suitable for computation which is converting to Numpy array. Then the processed image is fed into pre trained CNN model which is MobileNetV2 to detect whether the person is wearing mask or not .The system makes green bounding box if wearing mask and red if not wearing.

The entire face detection and mask classification happens in this model.

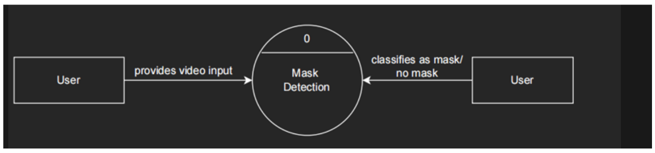

Fig. 2. DFD Level 0

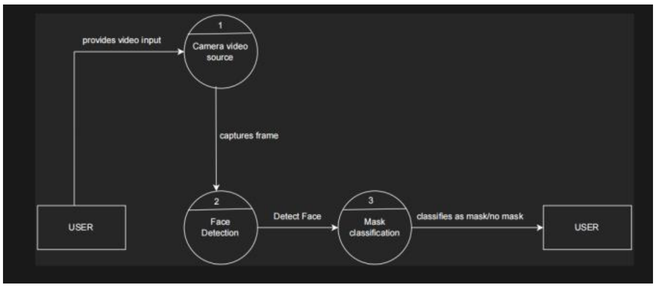

Fig. 3. DFD Level 1

Figure 2 and 3 shows the DFD Level 0 and DFD Level 1 respectively . DFD Level shows the the user gives video stream as input which comes from web cam .The system processes face and then determines whether each face has mask or not.

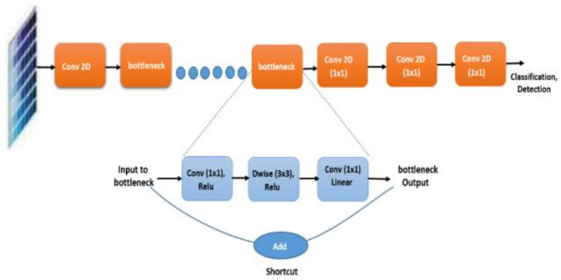

Fig.4. Architecture of MobileNetV2

Figure 4 shows the architecture of MobileNetV2 .The first layer is the convolutional layer which takes input as image and extracts features from the image then it consists of various bottleneck layers to reduce computational costs and extract more features and then last convolutional layer to classify images .

IV. PROPOSED METHODOLOGY

Dataset

The dataset comprises of 1915 images of people wearing mask and 1918 people not wearing mask

TABLE I.DATASET STATISTICS

Label	Image Dimensions	Total Number of Mask Images
With Mask	224x224	1915
Without Mask	224x224	1918

Data Pre -Processing

Before training data undergoes pre- processing images are resized into 224*224 pixels and converted into numpy array for easy computation. Text labels are also converted into binary labels using one-hot encoding.

Data Splitting

Data is split into training set (80%) and testing (20%).

Model Development

Data augmentation is done to diverse the dataset. Then it loads pre trained MobileNetV2 model without fully connected layer then fully connected layers are added to classify images as mask and without mask. The model is trained on dataset for 20 epochs. Before training data undergoes pre- processing images are resized into 224*224 pixels and converted into numpy array for easy computation. Text labels are also converted into binary labels using one-hot encoding.

Model Evaluation and Testing

Accuracy increases and loss decreases as the training proceeds and stops after gaining stability to avoid any overfitting. The prediction and confidence score is displayed in the video frame.

V. RESULT

These are the results that we get for our real time face detection model .

```
Epoch 6/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 14s 142ms/step - accuracy: 1.0000 - loss: 0.0413 - val_accuracy: 0.9909 - val_loss: 0.0569
Epoch 7/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 96s 1s/step - accuracy: 0.9776 - loss: 0.0826 - val_accuracy: 0.9935 - val_loss: 0.0459
Epoch 8/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 14s 143ms/step - accuracy: 0.9688 - loss: 0.1192 - val_accuracy: 0.9935 - val_loss: 0.0455
Epoch 9/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 98s 1s/step - accuracy: 0.9782 - loss: 0.0691 - val_accuracy: 0.9935 - val_loss: 0.0402
Epoch 10/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 14s 147ms/step - accuracy: 1.0000 - loss: 0.0312 - val_accuracy: 0.9935 - val_loss: 0.0402
Epoch 11/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 99s 1s/step - accuracy: 0.9824 - loss: 0.0631 - val_accuracy: 0.9922 - val_loss: 0.0402
Epoch 12/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 15s 154ms/step - accuracy: 1.0000 - loss: 0.0100 - val_accuracy: 0.9922 - val_loss: 0.0398
Epoch 13/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 98s 1s/step - accuracy: 0.9875 - loss: 0.0504 - val_accuracy: 0.9922 - val_loss: 0.0372
Epoch 14/20
95/95 ━━━━━━━━━━━━━━━━━━━━ 14s 145ms/step - accuracy: 1.0000 - loss: 0.0234 - val_accuracy: 0.9909 - val_loss: 0.0376
```

Fig.5. Iteration to check the loss and accuracy

	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

Fig.6. Model Evaluation

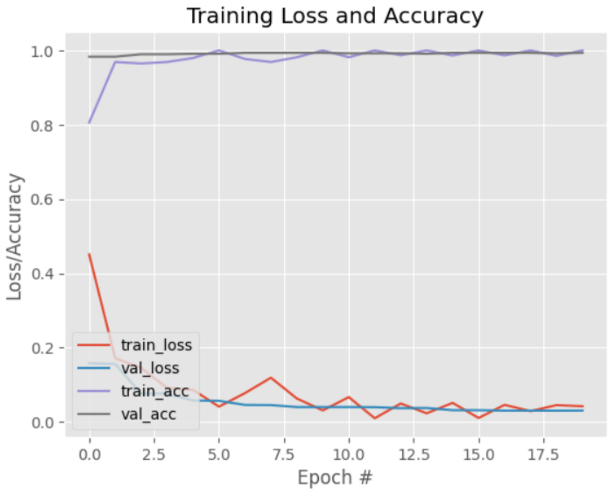

Fig.7. Graph of Training Loss and Accuracy

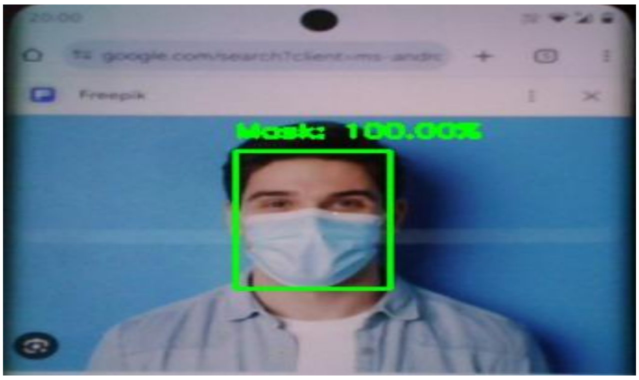

Fig.8. Result of a person wearing mask

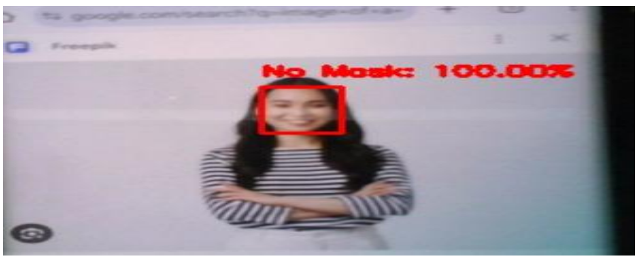

Fig.9. Result of a person not awering mask

Figure 5 shows the number of iterations on the datset which is 20 epochs . figure 6 is the model evaluation which shows an accuracy of 99 percent. Figure 7 shows Training loss and accuracy graph . It depicts the model is performing well with high accuracy and low loss. Figure 8 shows the image of a person wearing mask and figure 9 shows the image of a person not wearing mask .

VI. APPLICATIONS

Face mask detection has various applications especially in public sectors like hospitals ,schools , malls , colleges and etc. Real time Face detection helps to detect whether people are wearing mask or not. It became necessary to wear mask after the spread of coronavirus and after this we incident we really do not know what pandemic will occur next. With these types of pandemic it is obvious that they spread easily so to stop the spread of such viruses it becomes necessary to wear mask. In traffic signals it is not possible to check whether each person is wearing mask or not manually .So, we need a system or a model which detects this . These type of model plays a huge role in areas where there is huge amount of crowd like schools , colleges , offices and other public spaces.

VII. CONCLUSION

With the rise of Covid -19 we see the need to have a model that whithout any human intervention can detect if people are wearing masks in public places or not. So for this a deep learning model is created using CNN and Computer Vision. CNN uses light weight architecture which is MobileNetV2.It ia basically used to detect whether the person is wearing mask or not whereas OpenCV takes input from webcam and is used to localize the faces.MobileNetv2 is pre trained so it acts as feature extractor. It has various layers. It has Bottle Neck Layer which helps in computation while maintaining the accuracy . 1*1 layer uses ReLu activation function and it helps to extract more features but enhancing the depth then it has 3*3 depthwise convolutional layer , it is used to extract necessary features and then another layer lowers back the dimension . Then further layers are added on top of bottleneck layer to classify image into mask or no mask. Adam optimizer is used. Model runs 20 times over the entire dataset .Face detection is done by OpenCV and then live results are displayed on video stream . Green bounding box surrounds the face mask wearing mask and red bounding mask surround the face not wearing mask. It is displayed along with the probability score. Our model achieves an accuracy of 99 percent . This ensures that our model can be deployed in public places.

VIII. REFERENCES

[1] C.Jagadeeswari, M.Uday Theja, Performance Evaluation of Intelligent Face Mask Detection System with various Deep Learning Classifiers,2020

[2] Wei Bu, Jiangjian Xiao, Chuanhong Zhou, Minmin Yang, Chengbin Peng,” A Cascade Framework for Masked Face Detection”, 2017

[3] Amit Chavda, Jason Dsouza, Sumeet Badgujar, Ankit Damani ,” Multi-Stage CNN Architecture for Face Mask Detection “,2020.

[4] Aniruddha Srinivas Joshi, Shreyas Srinivas Joshi, Goutham Kanahasabai, Rudraksh Kapil, Savyasachi Gupta," Deep Learning Framework to Detect Face Masks from Video Footage",2020

[5] Reny Jose," A Convolutional Neural Network (Cnn) Approach to Detect Face Using Tensorflow And Keras",2019

[6] Adithya K1, Jismi Babu2," A Review on Face Mask Detection using Convolutional Neural Network", 2019

[7] "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network",2019

[8] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, Heling Chen, Yu Miao, Zhibing Huang, and Jinbi Liang," Masked Face Recognition Dataset and Application",2020

[9] Mohamed Loey, Gunasekaran Manogaran b,c , Mohamed Hamed N. Taha d ,Nour Eldeen M. Khalifa,," Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection",2020

[10] He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[11] A. Kumar, A. Kaur, and M. Kumar," Face detection techniques: a review,"Artificial IntelligenceReview,vol. 52,no. 2, pp. 927– 948, 2019.D.-H. Lee, K.- L.Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu, "Deep learning and control algorithms of direct perception for autonomous driving,2019

[12] Y. Fang, Y. Nie, and M. Penny, Transmission dynamics of the covid-19 outbreak and effectiveness of governmentinterventions: A data-driven analysis,"Journal of medical virology, vol. 92,no. 6, pp. 645–659, 2020.

[13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 4510- 4520, doi: 10.1109/CVPR.2018.00474.

[14] Xin, M., Wang, Y. Research on image classification model based on deep convolution neural network. J Image Video Proc. 2019, 40 (2019).

[15] Image Classification Using Convolutional Neural Networks Deepika Jaswal, Sowmya.V, K.P.Soman

[16] Deng, J., Guo, J., Ververas, E., Kotsia, I., and Zafeiriou, S., RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild, 2020, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 5202-5211, doi: 10.1109/CVPR42600.2020.00525

[17] He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[18] Bosheng Qin and Dongxiao Li, Identifying Facemask-wearing condition using Image super-resolution with classification network to prevent COVID-19, Sensors 2020

[19] Adam: A Method for Stochastic Optimization - Diederik P. Kingma, Jimmy Ba.