Real Time Face Mask Detector using Machine Learning, Python, OpenCV and Keras

Aniket Mishra
Student
Department of Computing Technology
SRM Institute of Science and Technology,
Kattankulathur , India

M.S. Amutha
Associate Professor
Department of Computing Technology
SRM Institute of Science and Technology,
Kattankulathur , India

ABSTRACT
After the new Coronavirus disease (COVID-19) case spread rapidly in Wuhan-China in December 2019, World Health Organization (WHO) confirmed that this is a dangerous virus which can be spreading from humans to humans through droplets and airborne.

As for the prevention, wearing a face mask is essentials while going outside or meeting to others However, some irresponsible people refuse to wear face mask with so many excuses. Moreover, developing the face mask detector is very crucial in this case.

This project aims to develop the face mask detector which is able to detect any kinds of face mask. In order to detect the face mask, a YOLO V4 deep learning has been chosen as the mask detection algorithm.

Keywords
Covid-19 Pandemics, Face Recognition, Detectors, Real Time systems, Production Facilities, Classification algorithms

1. INTRODUCTION
The spread of COVID-19 is increasingly worrying for everyone in the world. This virus can be affected from human to human through the droplets and airborne. According to the instruction from WHO, to reduce the spread of COVID-19, every people need to wear face mask, do social distancing, evade the crowd area and also always maintain the immune system. Therefore, to protect each other, every person should wear the face mask properly when they are in outdoor. However, most of selfish people won't wear the face mask properly with so many reasons.

To overcome this situation, a robust face mask detection needs to be developed. In order to detect a face mask, the object detection algorithm can be implemented. The state of art of object detection algorithm which has a robust performance is the You Only Look Once (YOLO). As presented in [1], Susanto, et al., used the YOLO deep learning method to distinguish the white ball and goal which is integrated to humanoid robot soccer.

In this paper, we proposed a real-time face mask detection system using YOLOv3 algorithm. In order to increase the processing time and accuracy of the detector, we employ Haar cascade detector to detect the face region in the input images, and then put the region of interest (ROI) into the YOLOv3 to detect the face mask. The deep learning model has been trained with dataset of 7,000 samples. Finally, we build up a whole system to demonstrate an application in which people are checked whether they are wearing face mask or not at the entrance door.

2. BACKGROUND
YOLO is a state-of-art algorithm for real-time object detection by Joseph Redmon et al. The authors released three versions: YOLOvl in 2016 [6], YOLOv2 in 2017 [7], and YOLOv3 in 2018 [8]. YOLO algorithm has been proved to be effective in many object detection applications such as vehicle detection [9], aerial target detection [10], pedestrian detection [11], etc. The feature extractor of the YOLOv3 contains 53 convolutional layers, and thus it is named Darknet-53. This is a hybrid approach between the previous network v2 and v1. The detail description of Darknet-53 is shown in Table I.

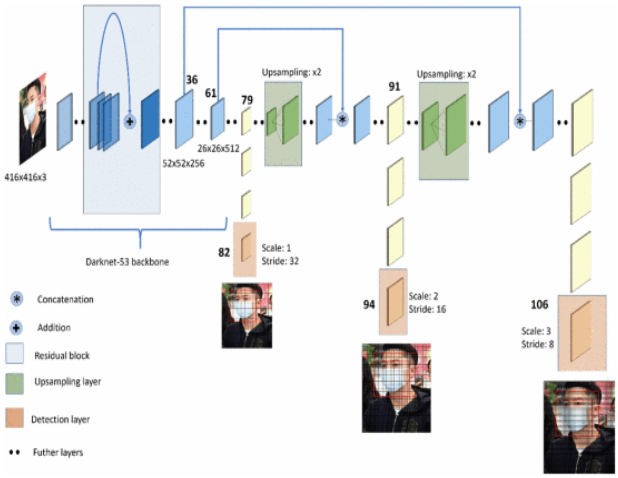

Fig. 1: Yolov3 network architecture

In this paper, we employ YOLOv3 to take its advantages for high accuracy and real-time processing.

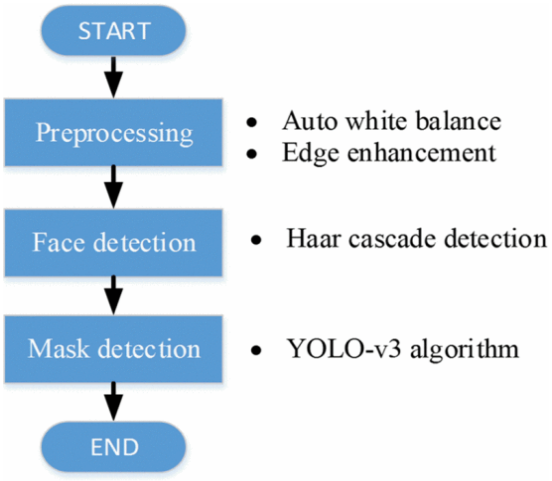

Fig. 2. Flowchart of Proposed Algorithm

3. PROPOSED ALGORITHM

The proposed algorithm for face mask detection consists of three steps: preprocessing, face detection, mask detection as shown in Fig. 2.

The preprocessing step is to enhance the input image quality by using auto white balance, and edge enhancement using unsharp filter. The auto white balance [12] is to ensure the color consistency of the input image frames in a various range of color temperature. The unsharp filter [13] is then applied to enhance the edges in the input images. Some researchers have proved that image enhancement can improve the accuracy of the object detection about 2-5% [14]–[15].

The face detection step is to detect the face region. We utilize the Haar cascade classifier proposed by Viola-Jones [3] to detect the face region. This classifier performs feature extraction by Haar Wavelet technique with 24×24 window size, uses AdaBoost to remove redundant features, and applies cascade classifiers to detect objects. The detected face regions by Haar cascade classifier is then put into the input data of the YOLOv3 algorithm to detect the regions of face masks.

4. LITERATURE SURVEY

Viola Jones Detector proposed a real time object model used to detect different classes of objects. It uses 24x24 base window size to evaluate any image with edge ,line and four rectangular features.Harr-like features are like convolutions to check weather given feature is available in the image or not This model fail to work in when image brightness varies even it exhibits poor performance when images are in different orientations.It has 97.5% accuracy with low complexity Methodology Used Convolutional Neural Networks (CNN) classification

Google Net architecture fundamentally uses inception method by constructing small convolution layers to reduce number of parameters it has around 22 layers of convolution and max pooling it is able to work effectively over Alexnet In this paper Deepnueral networks which adopts residual learning to train the models more deeper around 152 layers are used in this which is 8 times more than VGGnet This approach achieved relatively better performance in object detection over COCO data set but still has high complexity It has High accuracy with minimum complexity Methodology Used CNN

UNet and SEnet are used to perform segmentation of heart ventricular segmentation This model arrange's the weights in such a way that more weights are given to useful features and less weights are given to unimportant features Support vector machines are used to perform classification on objects which are going to build an equation for constructing line and classifies the objects based on the values mapped to this line.Shortcomings of Paper Time complexity achieved higher accuracy Methodology Used Semantic segmentation

VGG FCN

Tumuluru,Lakshmi Ramaniet. al. used CNN model to detect human face which is used efficiently in security related applications. They have collected various facial features such as mouth, nose, eyes stored as facial template and used it for detecting difference between faces.

Shortcomings of Paper time complexity. Achieved higher accuracy Methodology Used

Semantic segmentation VGG FCN

Tumuluru,Lakshmi Ramaniet. al. used CNN model to detect human face which is used efficiently in security related applications They have collected various facial features such as mouth, nose, eyes stored as facial template and used it for detecting difference between faces. Shortcomings of Paper Less data set Accuracy was less compared to training phase Methodology Used CNN FCN

Malathi, J. et. al., [14] mainly focused on identifying forgery images used in different places like in social media, and other publicity required places. Various techniques are proposed to find out features of a forgery image like image spicing,copy move attack which can be handled by using correlation analysis to find duplicate features. Shortcomings of Paper Lots of data duplication and redundancy low complexity with 97.85% Accuracy Methodology Used Corelation Analysis

Patelet. al., [15]proposed a model to find out the quality of the ironore by extracting the features from sample material in the mining industry It is very important to asses the quality of the ore . SVR support vector regressor used for online measure of the quality of the ore. In this process they have extracted 280 features extracted for object identification, SFFS was model was developed using SVR. Shortcomings of Paper Works for object detection but not for image analysis High accuracy in the case of object detection Methodology Used wavelet based nueral network

5. PERFORMANCE ANALYSIS AND EXPERIMENTAL RESULTS

Dataset

For the experiment, we use the datasetMAFA [16], which consists of 35,806 masked faces with a minimum size of 32×32. The face in this dataset have different orientation and occlusion degree. We select 7000 images which contain frontal faces from MAFA. The dataset is divided into 3 parts for training and validation and test set with 5000, 1000, and 1000 images, respectively. Fig. 1 and Fig. 3 show example images in the MAFA dataset.Units

Fig. 3. Images with people wearing mask

Fig. 4. Images with people not wearing mask or wearing mask the wrong way

We evaluate the performance of the proposed algorithm by two metrics: precision and recall, given by (2). Precision metric indicates the accuracy of the algorithm based on classification result. The recall metric presents the ability to find all relevant objects in a dataset.

Precision=True Positive/(True Positive+False Positive)

Recall=True Positive/(True Positive+False Negative)..(2)

6. RESULT AND DISCUSSION

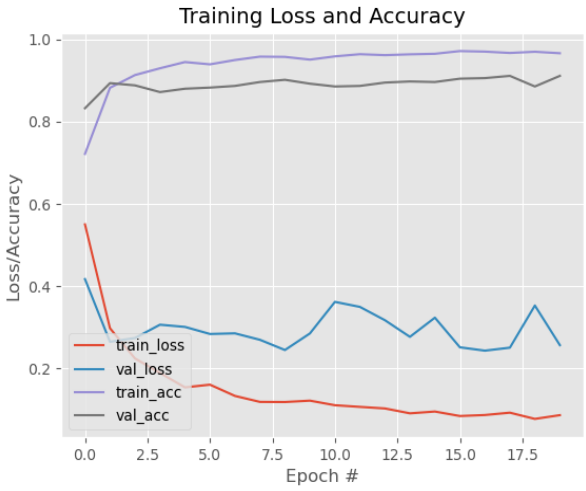

Fig. 4. Training Loss and Accuracy

This paper has presented a face mask detection system which uses the YOLOv3 algorithm and Haar cascade classifier. The proposed algorithm employs image enhancement technique to improve the accuracy of the system. Thanks to the advantages of the YOLOv3 network, the system can work in real-time with 30fps. This system can be applied effectively for practical applications to reduce the spread of infectious diseases such as Covid-19.

7. ACKNOWLEDGEMENT

I would like to express my special thanks to my guide M S Amutha who gave me the best guidance and helped me complete this journal on Real Time face mask Detector. She also helped me in completing my paper. I came to know about so many new things when working on this project. I am really thankful to them.

Secondly I would like to thank my parents and friends who helped me a lot in finalizing this project and also finishing it in the given time frame.

8. REFERENCES

[1] "World Health Organization (WHO)", *Coronavirus disease 2019 (covid-19): situation report*, vol. 96, 2020..

[2] "World Health Organization (WHO)", *Advice on the use of masks in the context of COVID-19*, June 2020.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition*, 2001.

[4] Wei Zhu, Qian-Liang Fu and Jun-Qi Bai, "The real-time object detection algorithm based on ORBP and cascade SVM", *2016 IEEE Advanced Information Management Communicates Electronic and Automation Control Conference (IMCEC)*, 2016..

[5] Hong Huo and Tao Fang, "Integration of bottom-up and top-down cues in Bayesian network for object detection", *Proceedings 2013 International Conference on Mechatronic Sciences Electric Engineering and Computer (MEC)*, 2013.

[6] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified real-time object detection", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[7] J. Redmon and Ali Farhadi, "YOLO9000: Better Faster Stronger", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[8] J. Redmon and Ali Farhadi, *YOLOv3: An Incremental Improvement*, 2018.

[9] Yunxiang Liu, Guoqing Zhang and Yuanyuan Zhang, "Vehicle detection method based on ADE-YOLOV3 algorithm", *International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, 2019.

[10] Lecheng Ouyang and Huali Wang, "Aerial Target Detection Based on the Improved YOLOv3 Algorithm", *2019 6th International Conference on Systems and Informatics (ICSAI)*, 2019.

[11] Changhao Piao, Xianhao Wang and Mingjie Liu, "Pedestrian Detection Using Optimized YOLOv3 in UAV Scenario", *2019 International Conference on Intelligent Computing Automation and Systems (ICICAS)*, 2019.

[12] Jintao Jiang, Ming Yang, XinPo Wang and Zhengguan Wu, "Auto White Balance Algorithm Based on Digital Camera", *2011 International Conference on Internet Technology and Applications*, 2011.

[13] R. C. Gonzalez and R. E. Woods, Digital Image Processing, AddisonWesley, Publishing Company, Inc., 1992.

[14] X. Liu, M. Pedersen, C. Charrier and P. Bours, "Can image quality enhancement methods improve the performance of biometric systems for degraded face images?", *Colour and Visual Computing Symposium (CVCS)*, 2018.

[15] M. O. Oloyede, G. P. Hancke and H. C. Myburgh, "Improving face recognition systems using a new image enhancement technique hybrid features and the convolutional neural network", *IEEE Access*, vol. 6, pp. 75181-75191, 2018.

[16] Shiming Ge, Jia Li, Qiting Yel and Zhao Luo, "Detecting Masked Faces in the Wild with LLE-CNN s", *IEEE Conference on Computer Vision* And pattern recognition, 2017.