Face Mask Detection System

Juhi Rajput
Student
ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Sonika Bhatnagar
Assistant Professor
ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Abstract

Coronavirus disease 2019 has affected and contrived the world seriously. One major protection and preservation method for people is to wear masks in public areas. About In the present scenario or storyline, there are no efficient and well organised face mask detection applications which are now in high demand for transportation means, densely populated areas, residential districts, large-scale manufacturers and other enterprises to ensure or guarantee safety. Moreover, many public service providers and contributors need customers to avail the service only if they wear masks correctly. Also this system could be used where we necessarily require face-mask detection for safety measures for example in current scenerio due to the Corona virus. This system should capture the face of person not wearing mask and breeching the rules and instructions and should send an alert to the respective officers for any further action against the person. This project can be integrated with systems for application in airports, stations, offices, schools, and public places to guarantee that public safety guidelines are followed.

Keywords: Disease, Coronavirus, Detector, Face Mask Detection, Safety

I. INTRODUCTION

In the COVID-19 crisis wearing masks is utterly obligatory for public well-being and controlling the spread of the pandemic. Face mask detection refers to identify whether a person wearing a mask or not and what is the location of the face[3]. The problem is closely related to general object detection to detect the classes of objects and face detection is to detect a particular class of objects, i.e. face[24].

Applications of object and face detection can be found in many areas, such as autonomous driving[13], education[14], surveillance and so on[11]. With the power of deep learning, I can develop a system that can identify whether a person has worn a mask or not and if any person is found breaching the rules it will send an alert to respective accountable officers. This type of projects could be integrated with lodged systems for application in airports, railway stations, offices, schools, transportation means, densely populated areas, residential districts, large-scale manufacturers and other enterprises to ensure that public safety guidelines and instructions are followed.

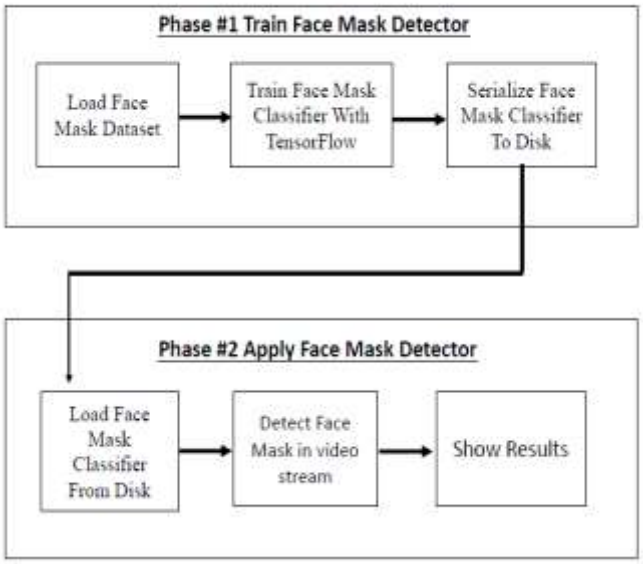
These systems therefore could be used in real-time applications which need face-mask detection for safety reasons due to the outbreak of Corona virus. Face detection is the main process for successive face-related applications like face recognition, facial expression recognition, and face mask detection, as its result straight influences on the successive application's performance. Therefore, face detection has become a research hotspot in the field of pattern identification and computer vision and has been widely studied in the past two decades[1].

At the moment, WHO recommends that people should wear face masks if they have respiratory symptoms, or they are taking care of the people with symptoms[19] . The likelihood of wearing face masks in public is increasing due to the COVID-19 epidemic all over the world. Moreover, numerous public service providers need customers to avail the service only if they wear masks[18]. Before Covid-19, People used to wear masks to safeguard their health from air pollution[2]. When it comes to identifying/detecting a face in still image or pictures and detecting a face in a real-time video stream, there is not much difference between them because videos are basically made up of frames, which are still images so we mounted the face detection on each and every frame in a video.

A. *Architecture of the Model:*

This project is two-stage Corona virus face mask detector, sketching how our computer vision/deep learning pipeline will be executed. The two phase face mask detector consists of phase one of training the detector and phase two of applying the mask detector by loading classifier and displaying results. I'll review the dataset I'll be using to train my custom face mask detector. I'll then implement a Python script to train a face mask detector on my dataset using TensorFlow. I'll use a Python script to train a face mask detector and review the results. Given the trained Corona virus face mask detector, I'll proceed to implement two more Python scripts used to detect face masks in real-time video streams. We will access video stream of any real time camera using IP address of camera and then detect faces without mask in video streams.

Two-phase face mask detector:



This system should capture the face of person not wearing mask and should send an alert to the respective accountable officers for any further action against the person.

B. Dataset:

The dataset covers a various images of masked faces or unmasked faces, including faces with masks, faces without masks, faces with and without masks in one image and confusing and baffling images without masks. As the face mask dataset is an appropriately small dataset where features might be difficult to detect, we used transfer learning to transfer the learned kernels and essence from networks trained for a similar face detection task on an extensive and large-scale dataset.

C. Transfer Learning:

In transfer learning we solve a machine learning problem and stores the knowledge obtained from it and focuses on applying it to a different but related problem.
In a nutshell, this system works by performing three main tasks

1) Object detection
2) Classification
3) Email Service

This paper is organized as follows. In Section II, we review related works on face detection and main approaches such as Feature Base Approach and Image Based Approach. The Object Detection methodology and algorithms are discussed in Section III. Section IV describes Classification Methodology, evaluation metrics, results, and ablation study. Finally, Section V deduces the paper and discusses the future work.

II. LITERATURE REVIEW

1) Mask R-CNN by Kaihan Lin, Huimin Zhao, Jujian Lv and Canyao Li was an better object detection model rooted on speedy and rapid R-CNN, has an majestic performance on various object detection and segmentation benchmarks and standards like COCO challenges and Cityscapes dataset[1].
2) Mohmed Loey and Gunasekaran Manogaran proposed a blended model utilizing deep and classical machine learning for face mask detection will be presented. The presented model consists of two components or stages in which first was to pull out fetures i.e. feature extraction and second was classification process of face masks from given frame as a final outcome[2].
3) Mingjie Jiang, Xinqi Fan and Hong Yan presented Retina Face Mask. Retina Face Mask was a single-stage face mask detector. It comprises of a feature pyramid network to blend high-level semantic details with various feature maps[3].
4) Detecting Masked Faces in the Wild with LLE-CNNs combines two pre-trained CNNs to extract candidate facial regions from the input image and represent them with high dimensional descriptors[4].
5) The purpose of this research by M Jannah, M Zarlis and H Mawengkang was to create facial expression system that captures image from the video camera. The system performance is evaluated by two parameters, detection rate and false positive rate[5].
6) In this paper, Ming-Yuan, Shieh and Tsung-Min suggested or put forward facial detection scheme, which is rooted on depth map analysis, focuses to enhance the effectiveness and fruitfulness of facial detection and recognition under different environmental illumination conditions[6].

7) In this paper Springer, Berlin and Heidelberg presented a speedy image pre-processing using the initiating of a straight shaded compact mask focusing over the faces[7].
8) Diego Socolinsky and Andrea Selinger in their research in Thermal face recognition in real life presented a comparative study of face recognition performance with visible and thermal infrared imagery[8].
9) In Facial emotion recognition using multi-modal information research it was concluded that Facial emotion recognition will become vitally important in future multi-cultural visual communication systems, for emotion translation between cultures, which may be considered analogous to speech translation[9].
10) Woody Bledsoe, Helen Chan Wolf, and Charles Bisson worked on recognizing human faces using the machine. Their early project was dubbed "man-machine" on facial recognition. In that project the coordinates of the facial features in a photographs had to be marked by a man then this photograpgh could be used by the computer for recognition. Takeo Kanade in 1970 openly revealed a face matching system that located material features such as the chin without human interuption and premeditated the facial features location. Later tests showed that the system could not always reliably and authenticaly identify or detect facial features. In 1993 the Defense Advanced Research Project Agency (DARPA) and the Army Research Laboratory (ARL) established or started the face recognition technology and developed systems that could be used or employed in a rich productive and high-yielding real life environment to assist and help security & intelligence[10].

III. RELATED WORK

A. Face Detection:

Feature Base Approach and Image Base Approach are the two main approaches for face detection. Unique features of object help in object recognition in Feature Base Approach. Statistical analysis and machine learning techniques are used to find the relevant characteristics of face and non-face images in Image-based Approach.

B. Feature Base Approach:

A human face consists of many features, which can be recognized between a face and many other objects. It locates and detects faces by extracting structural features like eyes, nose, mouth etc. and then uses them to detect a face. Generally, some sort of statistical classifier qualified then helpful to separate and distinguish between facial and non-facial regions. In addition, human faces have particular textures which can be used to differentiate and distinguish between a face and other objects. Moreover, the edge of features and traits can help to detect the objects from the face. By using OpenCV we will implement a feature-based approach.

C. Image Base Approach:

The learned characteristics are in the form of distribution models or attribute functions that is used for face detection. We will use different algorithms such as Neural-networks, HMM, SVM, AdaBoost learning in this section. By using MTCNN or Multi-Task Cascaded Convolutional Neural Network, which is an Image-based approach of face detection we will see how we can detect faces.

D. Object Detection:

Haar Cascade algorithm, also known as Voila-Jones[15] algorithm is used to identify faces. It is basically a machine learning object detection algorithm which is used to identify objects in a picture or video. It is a machine learning based approach where a cascade function is instructed or trained from a lot of positive and negative images. Then we can use this to detect objects in other images. This algorithm has four stages:

1) Haar Feature Selection
2) Creating Integral Images
3) Adaboost Training
4) Cascading Classifiers

In OpenCV, we have many trained Haar Cascade models which are saved as XML files We used this file instead of creating and training the model from scratch. We have used “haarcascade_frontalfacealt2.xml” file in this project. The pre-trained face detection model looks to work great for this case, and it detects faces even when they are partially or at a certain extent covered by masks. So, no need to re-train anything for the first task.

To detect the faces of varying sizes, this algorithm looks at subregions of the image in numerous scales. Now we have all the detections for the selected image. Detections are stored as pixel coordinates. Width and height of the rectangle that encompasses the detected face defines each detection by its top-left corner coordinates. We will draw a rectangle over it using OpenCV’s rectangle() draws rectangles over images to show the detected face.

E. Classification:

As there was no pre-trained classifier to distinguish or differentiate faces with and without masks, we need to develop a classifier that can differentiate between faces with masks and without masks. We need data in the form of images to create this classifier. We have a dataset accommodating image faces with mask and without a mask. So we finetune a pre-trained network called MobileNetV2 which is trained on the Imagenet dataset.

F. Methodology:

We took all the images and assigned them to some list and then classified them accordingly. Also, we preprocessed the image and resize it to 224x 224 measurements. We just removed the top layers of this pre-trained model and added few layers of our own to design it according to our problem. We have only two outputs so the last layer has two nodes. This is called transfer learning. Then we converted the labels into one-hot encoding. Then to evaluate them we will split the data into training and testing sets. The next step is data augmentation which remarkably increases the multiplicity of data available for training models, without actually collecting new data. Then we compiled the model and trained it on the augmented data.

G. Data Augmentation:

Techniques such as cropping, rotation, shearing and horizontal flipping are commonly used Data augmentation techniques to train large neural networks.

H. Evaluation Metrics:

We employed precision and recall as metrics, they are defined as follows.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Where TP denotes the true positive, FP denotes false positive and FT denotes false negative.

I. Email Service:

As discussed earlier, the system will capture the faces without masks and will send these images as an email to any official or responsible or accountable officer. For this purppose, this system will use Gmail SMTP server to send the email. For this purpose we will have to turn on the option of Allow less secure apps to connect with your Gmail account. To connect with the mail server python provides two classes in the smtplib module. We will use Context Managers to connect to the plain old SMTP server. These context managers automatically closes the connection at the end. Context managers ensure efficient resource handling and management. Our code becomes more readable and easy to maintain. Use the smtp.login() function to authenticate your account details.

IV. CONCLUSION

In this paper, we have proposed a novel face mask detector, which can possibly contribute to public healthcare and well-being. The architecture of this face mask detector consists of Haar Cascade algorithm for object detection and MobileNetV2 as the backbone. The backbone MobileNet can be used for high and low computation scenarios. We utilize transfer learning to adopt weights from a similar task face detection, which is trained on a very large dataset to extract more robust and powerfull features.

REFERENCES

[1] Kaihan Lin ,Huimin Zhao ,Jujian Lv, Canyao Li,Xiaoyong Liu, Rongjun Chen, Ruoyan Zhao, " Face Detection and Segmentation Based on Improved Mask R-CNN", Discrete Dynamics in Nature and Society, vol. 2020, Article ID 9242917, 11 Pages, 2020
[2] Loey M, Manogaran G, Taha MHN, Khalifa NEM. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement 2021;167:108288.doi:10.1016/j.measurement.2020.108288
[3] Jiang, Mingjie & Fan, Xinqi. (2020). RetinaMask: A Face Mask detector.
[4] Shiming Ge , Jia Li and Qiting Ye – "Detecting Masked Faces in the Wild with LLE-CNNs"- 2017 - IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 426-434
[5] M Jannah, M Zarlis and H Mawengkang – "Facial expression system on video using widrow hoff" – Journal of Physics: Conference Series, Volume 978, 2nd International Conference on Computing and Applied Informatics 2017 28–30 November 2017, Medan, Indonesia
[6] Ming-Yuan Shieh, Tsung-Min Hsieh, " Fast Facial Detection by Depth Map Analysis", Mathematical Problems in Engineering, vol. 2013, Article ID 694321,10 pages, 2013.
[7] Marco Grassi and Marcos Faundez-Zanuy- "Face Recognition with Facial mask Application and Neural Networks"- Series: Lecture Notes in Computer Science, Year: 2009, Volume 5398, Page 340 DOI: 10.1007/978-3-642-00525-1_34
[8] Diego A. Socolinsky, Andrea Selinger- "Thermal Face Recognition"- over time. 4. 187 - 190 Vol.4. 10.1109/ICPR.2004.1333735.

[9] De Silva, Liyanage & Miyasato, T. & Nakatsu , Ryohei. (1997). Facial Emotion Recognition Using Multi-Modal Information. Proceedings of the IEEE Intelligent Conf. Information, Comm. And Signal Processing. 1. 397 - 401 vol.1. 10.1109/ICICS.1997.647126.
[10] WoodyBledsoe , Helen Chan and Charles Bisson (1965)Computer to recognize human faces - Some Preliminary Results, Technical Report PRI 19A, Panoramic Research, Inc., Palo Alto, California.
[11] Z-QZ.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE transactions on neural networks and learning systems, vol. 30, no. 11, pp. 3212–3232, 2019.
[12] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," Artificial Intelligence Review, vol. 52, no. 2, pp. 927–948, 2019.
[13] 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," arXiv preprint arXiv:1910.12031, 2019.
[14] K. Savita, N. A. Hasbullah, S. M. Taib, A. I. Z. Abidin, and M. Muniandy, "How's the turnout to the class? a face detection system for universities," in 2018 IEEE Conference on e-Learning, e-Management and e-Services (IC3e).
[15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, vol. 1. IEEE, 2001, pp. I–I.
[16] P. Felzenszwalb, D. McAllester, and D. Ramanan , " A discriminatively trained, Multiscale , deformable part model , " in 2008 IEEE Conference on Computer Vision and Pattern Recognition . IEEE, 2008 , pp. 1–8.
[17] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," International journal of computer vision, vol. 128, no. 2, pp. 261–318, 2020.
[18] Y. Fang, Y. Nie, and M. Penny, "Transmission dynamics of the covid-19 outbreak and effectiveness of government interventions: A data-driven analysis," Journal of medical virology, vol. 92, no. 6, pp. 645–659, 2020.
[19] S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, "Rational use of face masks in the covid-19 pandemic," The Lancet Respiratory Medicine, 2020.
[20] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.
[22] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," arXiv preprint arXiv:1905.00641, 2019.
[23] R. Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.
[24] J. Deng,W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009, pp. 248–255.