International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2172122

Deep Learning Approach for detecting Covid-19 Face mask using YOLOv4 Algorithm

Rutuja R. Mahurkar[1], Prof. Naresh G. Gadge[2]
[1]Master of Technology Scholar, CSE, GCOE, Amravati, Amravati, Maharashtra, India
[2]Assistant Professor, Computer Science & Engineering, GCOE, Amravati, Amravati, Maharashtra, India

Article Info
Volume 7, Issue 2
Page Number: 650-654

Publication Issue :
March-April-2021

Article History
Accepted : 25 April 2021
Published : 30 April 2021

ABSTRACT

The Covid-19 is declared a pandemic all over the world by WHO on 11 March 2020. Various guidelines were issued by WHO for the prevention of coronavirus. One of the guidelines is wearing a face mask. From the various researches, it is proven that wearing a face mask minimizes the risk of virus transmission. Thus, a system is needed which reduces the load on governing body in the accomplishment of Covid-19 laws in crowded public places. A deep learning model using the YOLOv4 object detection algorithm is used for detecting whether people are wearing a mask or not, from images and video streams. In the proposed methodology, CSPDarknet53 is used for extracting facial mask features.

Keywords : COVID-19, WHO, YOLOv4, deep learning, CSPDarknet53, Face mask.

I. INTRODUCTION

Pandemic on global scale COVID-19 conditions arose amid a worldwide outbreak of dangerous disease. The situation is now under attack and is deteriorating rapidly in all of the WHO-designated countries. About 139 million people infected by coronavirus over all over the world and 2.99 million people lost their lives until now. Coronavirus affected almost every country in the world. The transmission of is virus is mainly through the air if the infected person exhales directly into the open air. Millions of individuals become ill in a single day itself. Throughout this crisis period, everyone can raise awareness and, of course, participate in some self-help activities. Through continuous evaluation and protest of people's wellbeing, the country's government, social authority, and workplace should strictly obey appropriate laws. Through handshaking, mouth germs, and exchanging accessories with others, this germ spreads virally in every area from one to many and many to millions [1].

The coronavirus outbreak has resulted in unprecedented levels of international scientific collaboration. With the help of machine learning and deep learning, researchers and clinicians are trying to develop various models for monitoring the spread of this virus and minimizing the infected population. This research will help the scientists for finding the proper medical solutions and preventive guidelines for reducing the spread of the virus and early detection of infections. The method of tracking large numbers of people, on the other hand, is becoming more complicated. Not every time human intervention is possible at crowded places for law maintenance. For this, we required object detection systems that can effectively work in real-time. There are many object detection algorithms are available. Like R CNN, Fast R CNN, YOLO, SSD, RetinaNet, etc. There are broadly two categories of object detectors i.e. regional based and classification based. Regional-based also called two-stage detection, as the name suggests detection performed in two steps. First, it detects the region of objects and then identifies the object. Some examples are R CNN, Fast R CNN, etc. Whereas Classification based also known as one-stage detectors can slide the window over the image and calculates the bounding box and corresponding class probabilities. YOLO family, SSD, Multi-box, G CNN is some of its examples [2]. The selection of object detectors depends upon the various factors like what speed we want, what accuracy we want, and available dataset, etc. Two-stage detectors achieve more accuracy than a single-stage but are relatively slower in real-time.

Hence deep learning approach i.e. YOLOv4 object detection algorithm can work well in real-time scenarios. Thus, the model is trained to detect facemask at public places to minimize the spread of coronavirus using the YOLOv4 object detection algorithm.

II. LITERATURE REVIEW

The deep learning approach has become quite popular in recent years. Almost every field, has its applications from medical fields to military security, etc. Below we have discussed how the different versions of YOLO are used for detecting various types of images.

L. Aziz et al. [2] survey 300 publications considering various advances in object detection. They explained two types of object detection procedures 1) regional-based like R CNN, Fast R CNN, etc. 2) classification-based SSD, YOLO, etc. They also considered 5 types of dataset. They analyze various factors affecting the prediction performance. Md. Rafiuzzaman Bhuiyan et al. [3] develop a facemask detection system with a YOLOv3 algorithm. The model is trained with both mask and no-mask images which consisting of a total of 650 images. The model achieved an accuracy of 96% for classification & detection accuracy. After 4000 epochs of training, the average loss is 0.0730. For training neural network CSPDarknet 53 framework is used. Detection of objects in still images is quite easy in comparison with video images, To tackle this problem B. Wang et al. [5] proposed a model for video object detection using Yolo. An experimental result shows that the model works well on ILSVRC 2016 VID dataset.

Automatic license plate recognition has been a popular topic of research. Its applications are traffic monitoring, speed control management & related to enforcement of traffic laws. R. Laroca et al. [6] develop a model that uses a Fast YOLO object detector for automatic detecting license plate recognition. The model achieved impressive results on the SSIG dataset. SSIG dataset contains 2000 frames from 101 vehicle videos. The proposed system gained a recognition rate of 93.53% and 47Fps. The model also shows good results on the UFPR-ALPR dataset. The dataset contains 150 videos and 4500 frames collected from the camera & also from the vehicles which are in motion, the thus proposed system achieved an accuracy rate of 78.33% & 35 Fps. Shubham Shinde [7] has developed a model, which can detect human actions from video streams using the YOLO algorithm. The model does not process the whole video but only certain frames after a specific interval of time. For each frame, it allocates an action label. Model is trained over LIRIS dataset and achieved an accuracy of 88.37%.

Dongqing Shen[8] proposed a fire flame prediction system using the YOLO algorithm and compared the results with other available deep learning algorithms, for finding the most effective algorithm. Some changes have to be made in the base YOLO algorithm so that, most optimized version of the YOLO algorithm can be implemented. The model achieves an accuracy of 70%. For increasing the steel strip, surface quality Jiangyun Li [9] proposed a model, which can detect defects over the surface of the steel strip, along with location as well as the size of the defected region. The model can predict six types of total defects. Then it trained over a customized dataset. The model obtained the detection of 99% with 83 FPS.

III. PROPOSED METHODOLOGY

In this section, we discuss the proposed methodology for training a custom facemask detector. Fig 2 shows the proposed methodology flowchart. Workflow is divided into 6 parts. The initial steps are data collection and data pre-processing. After that data augmentation and data annotation performed. Now the most important and crucial step has to be performed i.e. model training. Model is trained using the YOLOv4 algorithm. Lastly, testing was performed over real stream images and videos.

A. Data Collection

Data Collection is the first step in model training. The greater the data we have, the higher the accuracy we get. We required a greater amount of data with proper labeling for our model. For this, we assembled 900 images. All obtained images contain both mask and no mask images.

B. Data Pre-processing

Once data is collected, there are some images, which are not proper for our model. Therefore, we have to discard those images initially. We have 900 images each containing both mask and no mask images. Out of which only 800 are proper. Also before training, we have to resize each image in our dataset. For our model, we take it as 512×416. Therefore, the final dataset contains 800 images containing 400 each for both mask and no mask. Following fig1 shows some dataset images.

Figure 1. Dataset images

C. Data Augmentation

Data augmentation is necessary for getting better results. The more the large variety of data we have more will be the model accuracy. There are various data augmentation techniques available for image classification like flipping, cropping, rotation, and color modification, etc. But in our case data augmentation is directly performed by the YOLOv4 object detection algorithm. It uses various methods like CutMix, CutOut, MixUp, SAT and Mosaic data augmentation, etc.

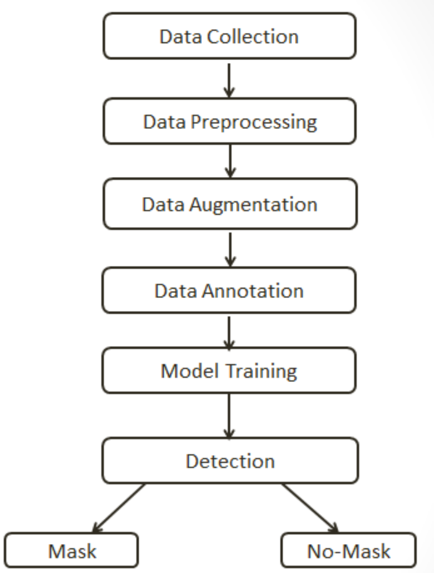

Figure 2. Proposed Methodology Flowchart

D. Data Annotation

The process of labeling the data refers to data annotation. To train deep learning-based model data needs to be properly label, so that model can quickly & accurately recognize input images. This also enhances the accuracy of the model. Data annotation is performed on images, videos, audios, text, etc. There are various tools are available for image annotation like Ybat, LabelImg, Labelbox, TrainingData.io, Hive, etc. Here Ybat[10] tool is used for the annotation of images in the proposed model.

E. Model Training

Recently, the YOLOv4 object detection algorithm was introduced in 2020. However, Joseph Redmon did not introduce it; according to him, there is the possibility that his innovations to be misused. Therefore, three founders Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao [4] found it.

YOLOv4 is faster and accurate than all its previous versions. YOLOv4 is 2 times as EfficientDet. It's AP and FPS improved by 10% and 12% to its previous version i.e. YOLOv3. In the YOLOv4 backbone, for feature extraction CSPDarknet53 is used. In its previous i.eYOLOv3, Darknet53 is used. DarkNet-53 is used in CSPDarknet53, a CNN and backbone for object detection. For the neck of YOLOv4 PANet path-aggregations used rather than Feature pyramid networks (FPN) as in YOLOv3. It has the same head as YOLOv3.

In the model, a training YOLOv4 object detection algorithm is used. First, the annotated image is passed into the model, where the YOLOv4 algorithm can divide the image into grid cells. Each cell separately predicts the object. Confidence score estimated for each grid cells along with its bounding box. The input image is divided into several grid cells, while each cell specifies the object features in that grid. For training CSP-Darknet53, the feature extractor acts as a network fundamental architecture. About 80% dataset is used model training. Once the model trains, then testing will be carried out over the remaining 20% dataset.

Figure 3. Model Workflow and detection

IV. CONCLUSION

A covid-19 global pandemic has triggered an immediate need in resolving a severe danger to the world. Hence, for minimizing the risk of covid-19 there is a need of wearing masks in public places. Thus, the proposed methodology can detect facemask with a deep learning model using YOLOv4 object detector architecture. YOLOv4 shows excellent results over real-time images and videos. This model helps in reducing human intervention and helps the governing bodies in proper Covid-19 law enforcement at crowded public places.

V. REFERENCES

[1]. S. Udhaya Kumar, D. Thirumal Kumar, B. Prabhu Christopher, and C. George Priya Doss"The Rise and Impact of COVID-19 in India", National Center for Biotechnology Information (NCBI), May 2020

[2]. Lubna aziz, Md. Sah bin haji salam (member, ieee), Usman ullah sheikh , and Sara ayub, Exploring Deep Learning-Based Architecture, Strategies, Applications and Current Trends in Generic Object Detection: A Comprehensive Review [2020] in IEEE Access Volume 8

[3]. Md. Rafiuzzaman Bhuiyan, Sharun Akter Khushbu, Md. Sanzidul Islam Dept. of SWE Daffodil International University Dhaka, Bangladesh, “A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3” (IEEE)2020
[4]. Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection”, arXiv:2004.0934 [cs.CV], 2020
[5]. Bin Wang, S. T.-B.-F.-D., “Detection and tracking based tubelet generation for video object detection. Journal of Visual Communication and Image Representation”, 102-111, 2019.
[6]. Rayson Laroca, E. S., “A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector”, 2018 International Joint Conference on Neural Networks (IJCNN), 2018.
[7]. Shubham Shinde, A.K.,“YOLO based Human Action Recognition and Localization”, International Conference on Robotics and Smart Manufacturing, pp. 831-838, 2018.
[8]. Dongqing Shen, X. C., “Flame detection Using deep learning”, 2018 4th International Conference Control, Automation and Robotics, IEEE, 2018.
[9]. First A. Jiangyun Li, S. B., “Real-time Detection of Steel Strip Surface Defects Based on Improved YOLO Detection Network”, IFACPapersOnLine, (pp. 76-81), 2018.
[10]. ”drainingsun/ybat”,GitHub. https://github.com/drainingsun/ybat.[Accessed: 20- Dec- 2020].

Cite this article as :

Rutuja R. Mahurkar, Prof. Naresh G. Gadge, "Deep Learning Approach for detecting Covid-19 Face mask using YOLOv4 Algorithm", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 2, pp. 650-654, March-April 2021. Available at doi : https://doi.org/10.32628/CSEIT2172122
Journal URL : https://ijsrcseit.com/CSEIT2172122