A CNN BASED SMART MASK DETECTION SYSTEM

Abhishek Gandhar[1*], Rajesh Holmukhe[2], Phagun Vijay[3], Madhav Sood[3] , Rudra Sharma[3], Adarsh Patra[3]

[1] Professor, EEE Department, EEE Department, Bharati Vidyapeeth's College of Engineering, New-Delhi, India,

[2]Associate Professor, Bharati Vidyapeeth (Deemed to be) University, College of Engineering Pune, India

[3]Undergraduate Scholar, EEE Department, Bharati Vidyapeeth's College of Engineering, New-Delhi, India

[1*]*abhishek.gandhar@bharatividyapeeth.edu*,

[2]rmholmukhe@bvucoep.edu.in,[3]*phagun.vijay@gmail.com*

ABSTRACT:

The Coronavirus made a new normalization of life where communal distancing and use of masks for covering their face perform an essential part in monitoring the effects of spreading of the corona virus, still the majority of population are found not using face shields or masks in public areas that accelerates the spreading of the corona virus. This might lead to the serious issue of rise in scattering of the disease. Therefore, to neglect any kind of circumstances we are in need to explore and alert the public for wearing masks. Persons can't be deployed for this procedure, as the risk of getting affected by corona virus increases. Henceforth, the presented model for mask detection is surrounded along the theories of artificial intelligence (AI), deep learning, object detection technologies and convolutional neural networks (CNN) which are the key subject of this project. The project performs by recognizing the people are wearing their face shields or masks or not in public areas via utilizing image processing and deep learning practices and transmitting data to the governing authorities. These algorithms for abject detection have been optimized for recognition of people with face masks or not. This paper is attempting for development of a model for real-time monitoring which will turn out to be pretty effective and simple. This model magnificently recognizes whether an individual is wearing a mask or not up to 98% of accuracy as achieved till date and observed that it has yielded outstanding outcomes for the detection.

KEYWORDS: Novel Coronavirus; CNN- Convolutional Neural Networks; AI -Artificial Intelligence ; Deep Learning ; Mask Detection .

DOI Number: 10.14704/NQ.2022.20.12.NQ77268 **NeuroQuantology2022;20(12): 2741-2751**

1. INTRODUCTION

On 11th March 2020, COVID-19 was proclaimed to be a pandemic by the World Health Organization (WHO). The COVID has already affected more than 20 million people and resulted in over 750 thousand deaths worldwide in accordance with the circumstances report 96 of the WHO. There are numerous severe large scale respiratory illness as well containing serious acute respiratory syndrome and Middle East Respiratory Syndrome. Such disorders

have grown a lot in the past few years. While certain countries, including Brazil, the United States and India are in continuous battle against this pandemic for a long time. India was under the toughest lockdown in its history to stop the coronavirus's spread.

Hence, people are continuously worried regarding the rapidly increasing varying health and respiratory disorders. Community health has been the topmost priority of the govt. of all the countries. The World Health Organization also stated that wearing face masks can facilitate to prevent coronavirus from spreading and the government made it mandatory for all the people to wear face masks while going out in public areas. The persons who are having respiratory syndromes should absolutely wear face masks. Hence, several service providers and super markets want all consumers to wear masks if they need their facilities. This is because face mask detection devices are essential to help people but there is quite a few researches concerning about face mask detection.

Face mask detection detects if a person is wearing a mask or not. It is similar to an object detection system in which it detects a specific class of items. By developing this system we are trying to assure people's protection in public areas. It can be put into operation in many places such as super markets and malls, colleges, schools, stations and so on. To achieve this task we will be utilizing one of the kind of major application of deep learning on computer vision known as the convolutional neural networks (CNN),designed in the 1980s by yann lecunn [1] , CNN technologies are a kind of neural arrangement which are systematically structured for arresting samples in multi-dimensionalspaces.

In 2015 , Joseph Redmon and his workmates launched an object detection system that executes all the necessary stages to determine an object by making use of a single neural arrangement for the first time, YOLO(You Only Look Once) algorithm[2]. It relabels the detection of object as a regression issue, straight from the image pixels to the coordinates of bounding boxes and probabilities of the given classes. This model continuously predicts many bounded boxes and probability of the classes for all those items covered by boxes. At its release time, YOLO algorithm has provided magnificent specifications that omitted the top-notch algorithms in terms of speed and accuracy for detecting object coordinates by utilizing the most recent version of YOLO(v5) for the mask detection that is framed in this research work.

2. LITERATURE REVIEW

Some of the notable contributions are summarized in this section.

a. In this image categorization by means of deep learning, the author Joseph Redmon et.al put forward "You Only Look Once" system for object detection in Real time[2].
b. Velantina et.al in 2020 , made a COVID-19 face mask detection

eISSN 1303-5150

model using the Caffe model[3].

c. Senthil Kumar et.al in 2017, reported the comparison of the two most widely used ma-Chine learning algorithms, K-Nearest Neighbor and Supporting Vectors Ma-Chine in his research work for face recognition [4].
d. Authors put forward enhanced YOLO (v5) a family of object detection framework and models already trained on the COCO dataset, representing an open-source research into upcoming vision of AI [5-6].
e. The author used kaggle depository to collect datasets for confirmation of exactness of pre-trained weights of YOLO [7].
f. Authors presented a clear understanding of a convolutional neural arrangement in this paper. The applications handles image data, such as very large image data set for image classification (Image Net), computer vision [8].
g. W. Zhiqiang and L. Jun (2017) described a review of convolutional neural arrangement based object detection technologies [9].
h. Authors gave voice about the most important condition for object detection refinement is to speed up the process. Depending on the general introduction to the context and the key solution Convolutional Neural Network [10].
i. A novel YOLO object detection model which is a CPU-based model planned to run on non-Graphics Processing Unit(GPU) computers is presented in this paper [11].

3. FACE MASK DETECTION ALGORITHM DEVELOPMENT

This segment of the paper provides a short overview of the methods came to use in the deployment of the detection technology .The method that has been developed comprises of predominantly four stages. The four stages are-data acquisition, data annotation, YOLO v5 model training, and lastly testing the model that has been developed. The methodology used in this project has been explained further along with the various libraries and tools required in the project.

3.1 The different libraries used here among other technologies:

1. NumPy:

NumPy stands for numerical python. NumPy is made up of matrix and a multi-dimensional array data formats. NumPy enables the user to perform mathematical operations on arrays.These mathematical operations may include statistical, algebraic or trigonometric operations depending upon the user.SciPy is developed using NumPy. SciPy in comparison to NumPy offers better performance and offers broader variety than NumPy. SciPy's better performance is key in performing various scientific and engineering applications. When an image is pre processed, it is enhanced or amplified to 224*224 pixels and is then transformed into a NumPy styled array. Post performing this process, precise labels are included in the dataset images.

2. TensorFlow

TensorFlow is used to perform high speed numerical calculations and was launched by Google on 9 November 2015. This is the most recent python library. TensorFlow is capable of developing deep learning models directly. Also, the wrapper libraries which help to create top of TensorFlow. This is one type of math library and machine learning applications like neural networks. There are many types of deep learning models available in the market and installation of TensorFlow is done a using pip command. TensorFlow helps in Data augmentation before it begins model training. It is also used to make algorithms prediction efficiency perfect, then download pre-trained image net weights. TensorFlow identifies whether a person is wearing a face mask or not using the camera installed on the laptop or it can also use a regular mobile phone camera for identification purposes. TensorFlow has a wide array of advantages, some of which are listed below:-

- Data augmentation.
- Pre-processing
- Loading Image data
- Classifier Loading.
- Make a completely new connected head

3. OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was constructed to give the users a common infrastructure for computer vision applications and to speed up the use of machine perception in the commercial products. Since OpenCV is BSD-licensed product,it is business friendly and is widely used.It offers various libraries which are used for various purposes such as object detection,face mask detection and deep learning algorithms.

3.2 Convolutional Neural Networks

CNN is one of the most important component of most deep learning based computer applications.

It was invented in the early 1980s by Yann LeCunn who was at that time a pioneer in the field of deep learning. CNNs are a type of neural network that are efficient at representing patterns in multidimensional spaces. This makes CNNs extremely handy in image processing, though they are used to process various types of data too. Every intricate neural network comprises of one or many *complex layers*, a software component that extracts meaningful values from the input image. And every complex layer is composed of various filters, square matrices that slide across the image and register the weighted sum of pixel values at different locations. Each filter has different values and extracts different features from the input image. The output of a complex layer is a set of "feature maps."

When stacked on top of each other, complex layers can detect a hierarchy of visual patterns. For example, the lower layers will produce feature maps for vertical and horizontal edges, corners, and other simple patterns. The next layers can detect more complex patterns such as grids and circles. As you move deeper into the network, the layers will detect complicated objects such as cars,trees and people.

CNNs primarily focus on the basis that the input will be made up of images. This directs the architecture to be set up in a way to best suit the need for dealing with the particular type of data.

eISSN 1303-5150

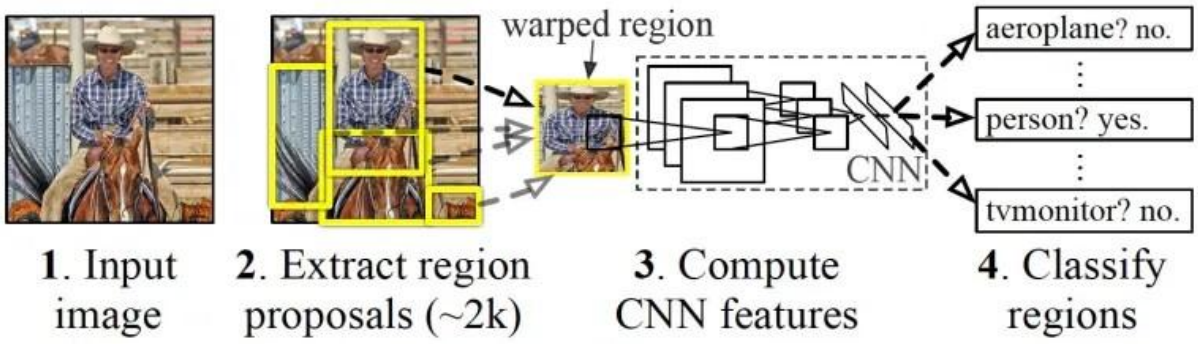

Fig.1. Process Flow Diagram of CNN

Through this simple method of transformation, CNNs are able to transform the original input layer by layer using convolutional techniques to produce class scores for classification and regression purposes.

3.3 YOLO - Object Detection Algorithm

Deep Learning contains a huge number of neural arrangement that utilizes multiple computer process cores and video processing cards to control a neural network neuron divided into a single node. It has its usage in many programs due to its famous nature. Here, YOLO study method is in use to detect mask.

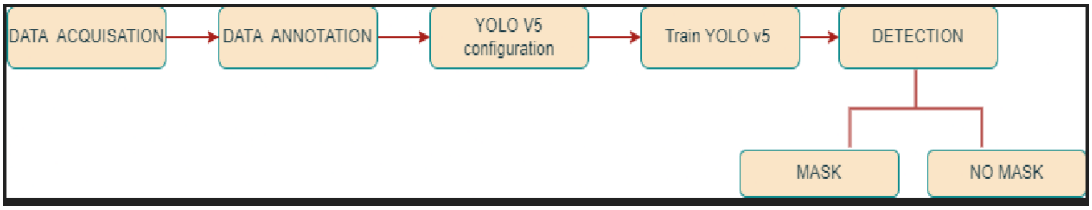

Fig.2. Outline of YOLOv5 algorithm

3.3.1 Benefits of YOLO

YOLO is in demand discovery algorithm due to its nature of attaining high accuracy while being able to work in real time. The algorithm "You only looks once" in the image means that it requires one forward transmission to the neural arrangement to make predictions. After non-max suppression, it provides known materials and binding boxes. At

eISSN 1303-5150

YOLO, one CNN simultaneously make predictions for multiple binding boxes and class opportunities in those boxes.

YOLO improves the performance of the detection as it trains with full images. YOLO has advantages over other methods of detection of the item : -

- In processing,YOLO is very fast.
- YOLO enters the code contextual information about classes and their appearance.
- YOLO learns the general representation of objects so that when trained in natural images and tested, the algorithm works much better compared to other advanced detection methods

3.3.2 Workflow of YOLO:

Workflow of YOLO will be discussed in detail under this section. Initially a small group of photographs is gathered. These photographs are used for the training purpose using YOLO. This group contains photographs of masked and non-masked people.

4. DATA ACQUISITION

Data is very essential for deep learning strategies. More the data we will use in AI training,the much better results we will get . In order to train YOLO, we require additional data and an appropriate notation. We use a web scraper tool to gather 900 images of masked people and non-masked people. These images cannot be put in use directly without them being pre-processed, so pre-processing of images is important before feeding into the model.

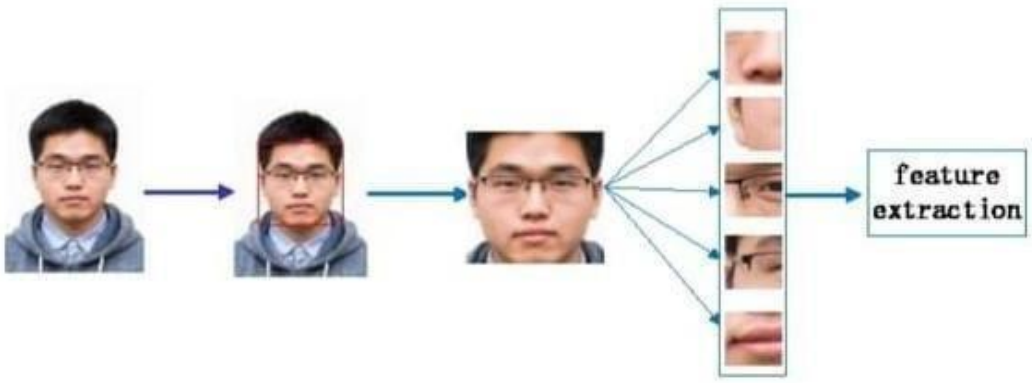

Fig.3.Feature Extraction Process

eISSN 1303-5150

5. DATA ANNOTATION

For the training process of YOLO we will require data annotation of images. For optimal working of the detection model, the dataset should be annotated well. A vast vareity of annotations are available; in this case, the "**bounding boxes"** technique is implemented. It forms a rectangle shaped margin over the dataset photographs. Annotation requires more time hence we will be using LabelIMG tool for data annotation.

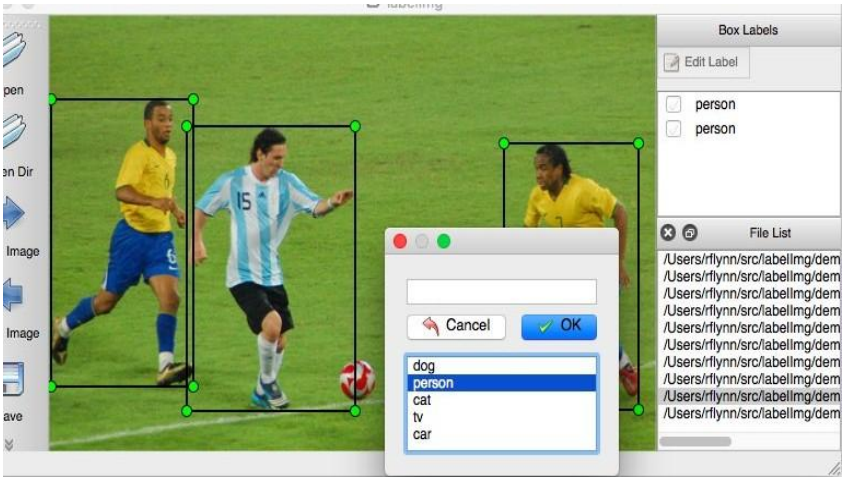

Fig.4 Image labeling by Labelimg

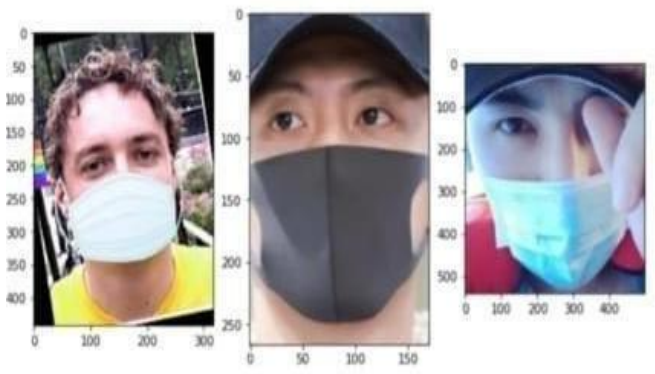

Fig.5 Sample Images of people wearing face masks

eISSN 1303-5150

6. MODEL TRAINING

The next step after setting up the model is to train it. The process of training requires a lot of time. The end outcome of the detection model totally rely on the quality of this procedure. In this system, we have a dataset which consists of all the information required for mask detection known as **'Pre trained weights'** that can be downloaded via yolo repository. They are then trained on our given data set of images . All the data is labeled using the help of a software tool. After the training process we get our own trained weight on the basis of our data provided for 2 classes segregated into 'mask' and 'no mask' categories which can be used in image and video testing.

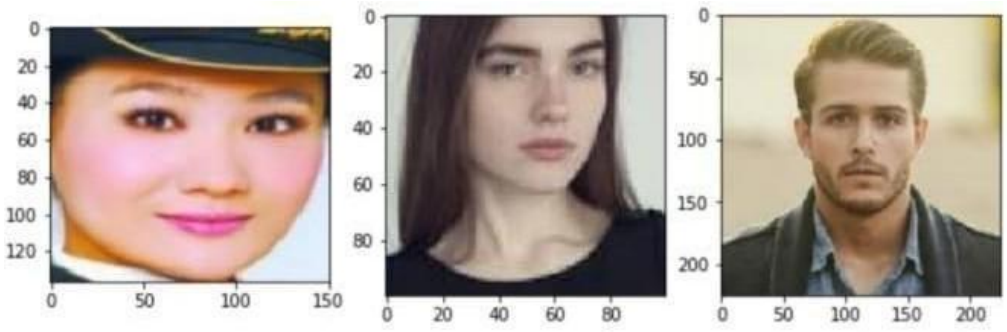

Fig.6 Model set-up for training data

7. RESULTS AND DISCUSSIONS

Now we will discuss the outcomes of our model. The model is examined via the dataset. The detectionmodel draws a comparison on trained weights and then the system categorizes the image into mask or no mask. The model has been trained into detecting mask and no mask cases during real time image capturing. To fulfill this purpose, we have used "OpenCV", using which the model can load the images for the testing process on the testing dataset provided by the user. Google Collab was used to complete the process. The model works accurately and can classify individuals wearing masks and not wearing masks with precision. We used photos from test library and using processed weights we got the following results :

Fig.7. Obtained Result Images

eISSN 1303-5150

The detection model accuracy was achieved around 98% for the given test datas. Also we tried to run our tests on live stream which turned out to be very accurate results as it saves the photos in a sequence and then after running the tests on them showcases the detected images.

8. CONCLUSION

A system has been developed which can be used in monitoring areas via real-time monitoring devices, without the need of any other device. The suggested model is a video analyzer which can be used for real time monitoring. It can be operated to detect wearing of masks and following protocol in public areas such as supermarkets, apartments,schools, colleges ,or other public places, where public gathering takes place. This can help in defeating the widespread of COVID-19 virus and is a powerful tool that can be used in precarious measurements. The system can be installed on the main entrance passage and can check the individuals who pass it. After processing the video in real time we can detect whether the person is wearing a facemask or not and prevent their entry in certain areas to minimize the spread of virus. If the individual is wearing his/her facemask, the entrance door will open; otherwise it will emit a warning prompt saying"kindly follow the protocols and wear a facemask". For the processing of images and real-time videos, the model uses the YOLOv5 and TensorFlow. The results can conclude that the developed model can be used in detecting facemask on an individual's face with high accuracy. The model quickly learns the parameters and collects the video from the camera, processes the video, identifies the objects, and concludes mask or no mask on the individual.

The studied model can also be enhanced through inclusion of numerous variables such as number of individuals, social distancing guidelines or could be further modified to be useful for the detection of helmets .

9. REFERENCES

1. LeCun, Yann, and Yoshua Bengio. "Convolutional networks for images, speech, and time series." (1995).
2. Joseph Redmon, S. D. (2016) “You Only Look Once(YOLO) Unified, Real Time Object Detection” IEEE.
3. Vinitha, Velantina, and V. Velantina. "COVID-19 facemask detection with deep learning and computer vision." *Int. Res. J. Eng. Technol* 7.8 (2020): 3127-3132.
4. Senthilkumar, Dr & Mano.MU, & Prakash.M, Saran & kan, Jaya. (2021). Face mask detection using AI and IOT , Erode Anna University : Chennai 600 025.

eISSN 1303-5150

10.13140/RG.2.2.23151.15521.

5. A. Patel, D. R. Kasat, S. Jain and V. M. Thakare, "Performance Analysis of Various Feature Detector and Descriptor for Real-Time Video based Face Tracking", International Journal of Computer Applications, vol. 93, no. 1, pp. 37-41, 2014. Available: 10.5120/16183-5415
6. S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi:10.1109/ICEngTechnol.2017.8308186.
7. Krizhevsky A., Sutskever I. and Hinton G. E. 2012 Image net classification with deep convolutional neural networks *Advances in neural information processing systems* 1097-1105.
8. Understanding of Object Detection Based on CNN Family and YOLO Juan Du1 Citation Juan Du 2018 *J. Phys.: Conf. Ser.* 1004 012029
9. W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," 2017 36th Chinese Control Conference (CCC), 2017, pp. 11104-11109, doi: 10.23919/ChiCC.2017.8029130.
10. Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9, 611–629 (2018).
11. M. Bansal, "Face Recognition Implementation On Raspberry Pi Using Opencv And Python", International Journal Of Computer Engineering & Technology

WEB URLs

1. https://github.com/thisisabusufyan/Real-Time-Face-Mask-Detection-using-YOLOv5-OpenCV-and-Python-Detect-mask-real-time-video-streams
2. https://pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorfl ow-and-deep-learning/
3. https://www.kaggle.com/datasets/ashishjangra27/face-mask-12k-images-dataset
4. Yolo v4 image detection

eISSN 1303-5150