International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 08 | Aug 2021 www.irjet.net p-ISSN: 2395-0072

FACE MASK DETECTION USING MACHINE LEARNING

RONAK BHOGAVALLI, B TECH, A[1]

[1]*Department of Mechatronics Engineering, SRM Institute of Science & Technology;*

ABSTRACT: *The COVID-19 pandemic has caused major crisis with respect to the health of most of the human beings in most of the places of the world which leaves a greater impact by affecting the health of the people. Major effective methods are made by maintain social distancing and compulsorily wearing a mask. By wearing a mask, it mainly reduces the risk of transmission of the disease. We are trying to present a hybrid model using classical machine learning algorithms, deep learning for the detection. The dataset includes images with and without mask where we try to make use of OpenCV for the real-time detection using a webcam. As it is compulsory to wear face masks in public areas and for enhanced safety of people, we make sure to implement such systems for safety and security reasons. The same model can also be use in workplaces to ensure all the employees wear their masks throughout the day. We are using dataset to build a COVID-19 face mask detector using computer vision, TensorFlow, python and Keras. Our main goal is to identify whether a person on image/video stream is wearing a face mask or not with the help of deep learning and deep learning. As face recognition also represents the modalities of biometric.*

1) INTRODUCTION

1.1 MACHINE LEARNING:

Machine Learning automates analytical model building knows as data analysis. It is a subset of one of the most popular intelligence known as Artificial Intelligence which is based of the data which systems can ideally learn from that is the preprocessed data then it identifies the patterns according to the previous data and it makes prior decision with minimal human intervention. It was eventually born from the pattern recognition and the theory that computers can easily learn without being programmed to perform some specific tasks as researchers are interested mainly in artificial intelligence and wanted to see if computers can learn from data. The main iterative aspect of machine learning is very important as it models re exposed to new data hence, they would be able to adapt independently. The main interest in machine learning is developed by some factors that have made Bayesian Analysis and data mining popular than ever. The highest and most available varieties of computational processing, data is much cheaper, but powerful and comes with an affordable data storage. Machine learning is mainly used in industries like finance, retail, healthcare, oil& gas, transportation etc. it is mainly used in those industries as it can successively predict and analyze the data eventually and better accuracy with the output data.

1.2 DEEP LEARNING

Deep learning is a learning that is similar to a learning and functions of the human brain. It is mainly also known as the subset of one of the popular technologies that is machine learning. It mainly works as similar to the human brain because this contains something known as neural networks in the working process but since it is manmade it is known as artificial neural networks. It mainly works without human intervention and it can also draw data from unlabeled and unstructured datasets. It uses a hierarchical standard of artificial neural networks to carry out the whole process of training which is used for machine learning algorithms. The artificial neural networks are built similar to the neuron present in the human brain and the main neuron nodes are connected together like a spider web. There are 3 types of neural networks used in deep learning which are as follows:

- Artificial Neural Networks (ANN): Figure 1.1: ANN It is a multiple group of neurons/perceptron's at each and every layer and it is also known as Forward Neural Network because all the inputs are processed only in the forward direction. It mainly consists of 3 layers that is Hidden, Input & Output. The input layer accepts the inputs from the system, the hidden layer processes the inputs from the systems and the output layer produces the result

,

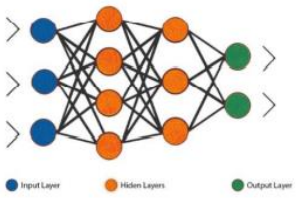


Figure 1.1: ANN

- Convolutional Neural Networks (CNN):

These types of neural networks are used mainly used in applications such as image and video processing projects and these neural networks play a very important role in the deep learning community. The main building blocks of these NN are kernels which are basically filters for these NN, they are mainly used to extract features from the input data using the

convolution operation. The main advantages of CNN are: that these NN filters automatically without mentioning it explicitly and they help in extracting the relevant and right features from the input data.

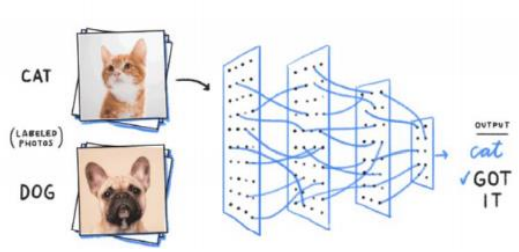


Figure 1.2 : CNN

- Recurrent Neural Network:

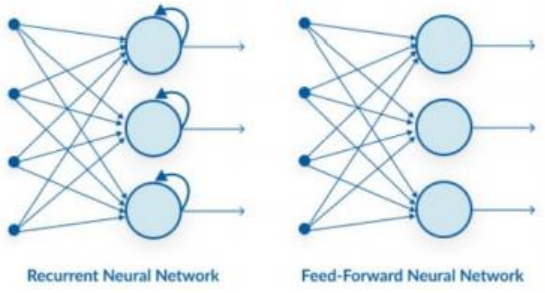


Figure 1.3: RNN

Recurrent Neural Networks are types of neural networks which are mainly used for voice recognition and for voice search etc. It is the one of the first algorithm which used by companies like Apple & Google's voice search for which the algorithm remembers its input because of an internal memory which makes it suited perfectly for problems dealing with sequential data in the field of machine learning

1.2 Computer Vision:

It is a sub-field of Artificial Intelligence that trains PC/Computers to understand and interpret the visual world. The input data in a computer vision system is done real time or by a preloaded image/video for which the computer then decodes and detects information or detects entities from the input data. It works very similar to a human vision except those human beings have a head start. As human sight has a very good advantage that is they contain the context to train and how to tell objects apart, how far away they are and if there is something wrong going on or not. As a human being gets trained since childhood with all the memories, he/she sees with the naked eyes and as this training leads to predicting the next step in a situation after being viewed from the naked eye. Computer Vision requires an input source that is a webcam or a CCTV camera.

Figure 1.4: Pedestrian Detection

Figure 1.5: Vehicle Detection

1.3 OPEN CV:

Open CV is basically an open-source computer vision library which is mainly used for solving computer vision problems in a program, in most of the cases the program is always written in Python.

It was mainly built for providing a common infrastructure for computer vision applications and to accelerate the use of machine learning in commercial products. This library has more than 2500 optimized algorithms which also includes a comprehensive set of both 5 classic and normalized computer vision and machine learning algorithms. These algorithms are mainly used to detect objects, faces, classify human action in videos and also tries to track the movements of objects in videos.

Figure 1.6: OPEN CV

1.4 TENSORFLOW:

It is mainly an open-source software which is used mainly to train and develop machine learning models and algorithms. It has a flexible, comprehensive

ecosystem of tools and libraries and community resources which enable a bright use of such libraries in the training models. The advantages are that models can be easily bult, we can easily deploy and train models in the cloud or on any device with whatever language you code in, it way too powerful and effective. Most of the people and industries who work with machine learning and Artificial intelligence use the TensorFlow library as it makes things easier for them.

Figure 1.7: TensorFlow

1.5 KERAS:

It is one of the most used deep learning frameworks by most of the users because it makes it easier for users to run new experiments and it empowers the users to try out new ideas and to increase their competition with respect to ideas and innovation. One main advantage is that the users can easily deploy capabilities of the TensorFlow platform and they can easily export models such as JavaScript to run directly in their web browsers

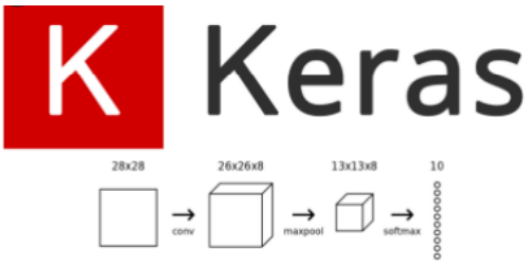


Figure 1.8: KERAS

1.7 APPLICATIONS:

In the current situation with the COVID-19 pandemic it is very important for every individual for wearing a mask in public areas as this reduces the transmission of the virus from person to person. Face Mask Detector being a tool which detects if a person is wearing a mask or not. This type of project can be used in most of the public areas where cameras can detect if individuals in public areas are wearing masks or not. It is an effective system that enables to check and detect if people are wearing masks or not, as it is done on a real time basis so the algorithms and codes should directly be implemented on the existing CCTV cameras.

Figure 1.10 Application of Face mask Detector

Figure 1.11 Application of Face Mask Detector

II PROPOSED SYSTEM

2.1 METHODOLOGY:

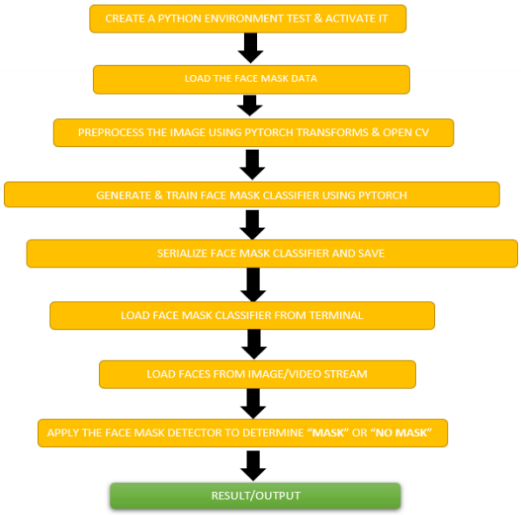


Figure 1.9: Methodology of the project

2.2 NEED OF THE PROJECT:

In this time during the COVID pandemic wearing of face masks have become really very important and crucial for people to wear masks and a project like this will help the government for surveillance, monitoring and safety purposes. As this be easily adapted with the CCTV

cameras so it is better that a system like this is gets implemented soon. Real time information can be identified, captured and be recorded for maintaining safety standards in a certain place. Instead of a person monitoring another person for wearing a mask which is very risky, it will be better if the system can identify the following aspect.

Figure 1.12 Need of the Project

II) WORKING & IMPLEMENTATION

2.1 Dataset:

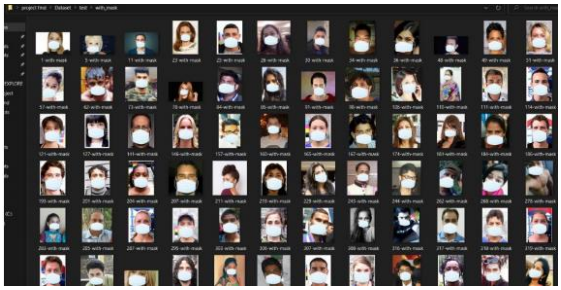


Figure 1.13 Dataset with Facemask

Figure 1.14 Dataset without Mask

The dataset was divided into 2 categories that is with mask and without mask so that the model can have better accuracy compared to the model which are just trained with just the face mask data and it is mainly used for the input dataset for training to detect for a real time video.

2.2 Data Preprocessing:

Most of the images have been augmented using open-source computer vision library (OPEN CV) which was then applied to all the input raw images to that those can be converted to clean versions which could easily be used as input for the neural network.

2.3 Extracting the images from video samples:

We cannot train the neural network in the video data; we need different and a huge number of images to train the model in order to train the machine. The video specifications are at 30 frames per second and all these frames get converted directly into images and then hereby result to a total of 60,000 images

2.4 Data Accretion:

In the first phase of Deep Learning training, it is very important to have a lot of data so that the model can learn all the variations in the images. A common method to increase the number of training sets is to use data accretion. It was also used to generate new images by performing a set of accretion operations on the images which were extracted from video frames.

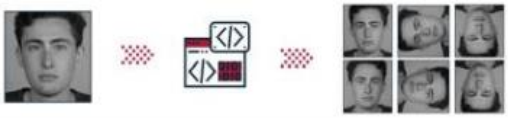


Figure 1.15: Data Accretion

2.5 Extraction Coordinated from Images

To locate and find main parts of the face such as mouth, nose, jawline, eyes facial coordinates are used. These coordinates are essential for head pose estimation, blink detection, yawning detection, etc. An open-source C++ library has a pre-written function that can be used to obtain facial coordinates. This library is programmed to find the y & x coordinates of more than 67 facial coordinates to map the facial structure.

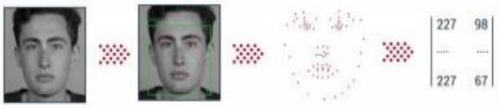


Figure 1.16: Extraction Coordinates from Images

2.6 Training the model:

While training we need to made a virtual python environment and name it 'test' and then activate it and then install all the requirements in it and then we can start the training by having an optimal number of epochs for the CNN and the training of the neural network is very necessary as the model learns patters that are specific to sample data to a great extent. The total number of epochs used in this project is 20.

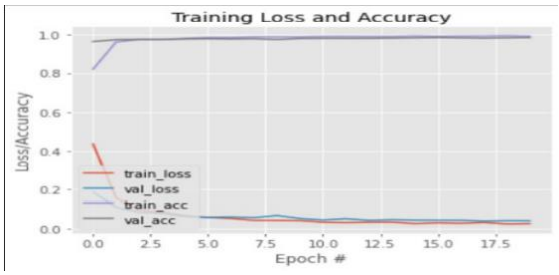


Figure 1.17: Training & Accuracy

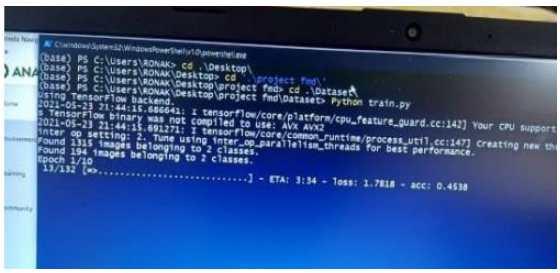


Figure 1.19: Training & epoch values for the model

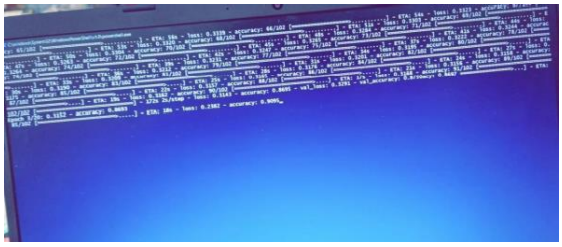


Figure 1.20: Training & epoch values for the model (1)

III) RESULT

3.1 Result & Output:

The output of the system detects if a person is wearing mask or not and it mainly also shows the accuracy. If a person is wearing a mask and depending of him/her wearing his/her mask an amount of accuracy is predicted by the system.

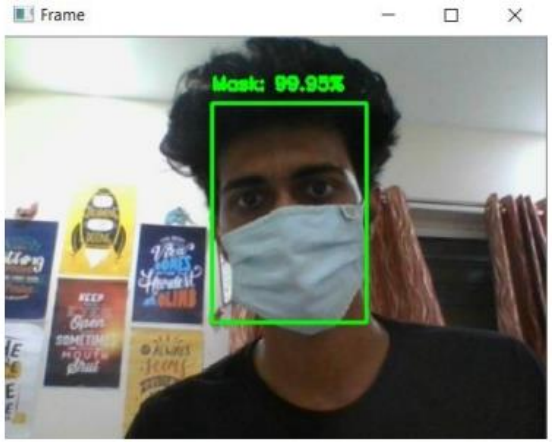


Figure 1.21: Output video for a single Individual

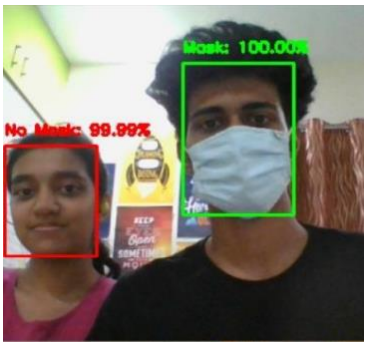
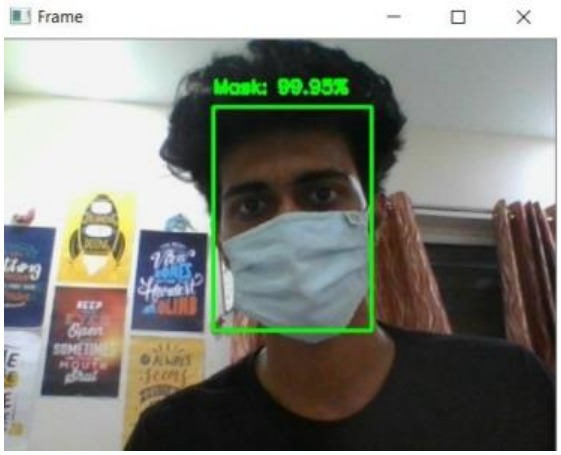


Figure 1.22: Output Video for many individuals.

The maximum accuracy of the model was 98% for face mask detection after training. The frameworks which were used are OpenCV, keras, TensorFlow, mobileNETV2 etc. This model is very accurate since the MobileNETV2 architecture was used it became very efficient and it was easy to deploy the model to embedded systems and also the face mask detector did not use any morphed masked images in the dataset. This type of system can be used in real-time applications like for corporate companies, schools, classes for maintain safety standards due to the outbreak of COVID-19.

The mask detection happened in 2 cases:

- **Identification on live webcam**: The system detects if person/persons is/are wearing a mask or not and it mainly detects it in a live video stream with an accuracy of 98%



- **Identification for an input image**:

  The system detects if the particular person is wearing a mask or not for a provided input image and detects identifies if that particular person is wearing a mask or not in that input image.

Figure 1.23: Input Image

Figure 1.24: Output Image after face mask detection.

IV) CONCLUSION & FUTURE SCOPE

4.1 Conclusion:

The project was completed in stipulated time and met all the objectives. As companies are emerging with high and innovative technologies, some of the technologies should also be used for maintaining enhanced safety for the people which is turn is a benefit for the society. The models were tested real time video-streams and images with an optimal accuracy and as the optimization of the system is a constant process and as it is accurate solution after considering the hyper parameters and a specific model could be used as a case for the edge analytics.

4.2 Future Scope:

In the near future we tend to integrate this model with the temperature sensors and blood pressure sensors by which if a person would stand in front of a screen the camera would detect if he/she is wearing a mask or not, the infrared sensor will the check the temperature of a person and hence the applications of each sensor would be used and the person will get an optimal report with the data on the screen.

V) REFERENCES

1) P. A. Rota, M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp,K. Maher, M.-h. Chenet al., "Characterization of a novel coronavirus associated with severe acute respiratory syndrome," science, vol. 300, no. 5624, pp. 1394–1399, 2003.

2) Z. A. Memish, A. I. Zumla, R. F. Al-Hakeem, A. A. Al Rabeeah, and G. M. Stephens, "Family cluster of middleeast respiratory syndrome coronavirus infections,"New England Journal of Medicine, vol. 368, no. 26, pp.2487–2494, 2013.

3) Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, "The reproductive number of covid-19 is higher comparedto sars coronavirus,"Journal of travel medicine, 2020.

4) 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.

5) N. H. Leung, D. K. Chu, E. Y. Shiu, K.-H. Chan, J. J. McDevitt, B. J. Hau, H.-L. Yen, Y. Li, D. KM, J. Ipet al.,"Respiratory virus shedding in exhaled breath and efficacy of face masks."

6) S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, "Rational use of face masks in the covid 19pandemic,"The Lancet Respiratory Medicine, 2020.

7) Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Peiet al., "Masked facerecognition dataset and application,"arXiv preprint arXiv:2003.09093, 2020.

8) Z.-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.

9) 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,20.