IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

Video Based Face Mask Detection and Face Recognition using CNN, YOLO and Google Facenet

Dhvani Fajalia[1], Rushikesh Satkar[2], Abhijeet Jejurkar[3], Prof. Sagar Dhanake[4]
Student, Computer Engineering[1, 2, 3,]
Assistant Professor, Computer Engineering[4]
D Y Patil Institute of Engineering and Technology, Ambi, Pune, Maharashtra, India

Abstract: *Globally Covid-19 has affected a large number of people. Covid-19 caused disruptions to many aspects of normal life. This also caused an increase in the death rate and Tension and Fear among the people. To avoid and overcome all these challenges WHO asked people to follow safety measures to avoid the spread of the virus. Wearing a mask is one of the safety measures among this. Mainly sometimes people gather at public places like offices, banks, airports, bus stops, etc. Considering this condition, we have designed a system that helps in mask detection in the various surveillance system at such places, also it helps in the face recognition of those people who have not worn the mask. Previous Systems were either only doing face recognition or face mask detection we came forward with this new approach by combining these two algorithms. All This is Possible by Using Artificial Intelligence which helps in Image Processing and Machine learning. Algorithms Like CNN, YOLO, and Google FaceNet were used for the same. Using this System, we got 97% Accuracy in Detecting a Face mask by using CNN. This System helps in classifying with mask and without mask person in the video and also help to identify the person who has not worn a mask which helps to send a notification to that person who violates the rule..*

Keywords: Convolutional Neural Network, Artificial Intelligence, Machine Learning, you only look once

I. INTRODUCTION

Do you know Globally, as of December 2021, there have been Cases 27.1 Crore confirmed cases of COVID-19, including 53.1 Lakh deaths, reported to WHO? Coronavirus disease is an epidemic in the current date that forced an international health emergency. In recent times, the dangers and fears of Coronavirus are still great. Covid19 is broad group of viruses that spread diseases that range from flu to deadly diseases like MERS (Middle East Respiratory Syndrome) that is viral respiratory illness that is new to humans and SARS (Severe Acute Respiratory Syndrome). Many people are unable to protect themselves and their environment from the pandemic. Coronavirus spreads mostly from person to person through airborne transmission, mainly through close contacts. WHO has come with various Preventive measures to avoid the Spread of Covid19.

1.1 Motivation

Wearing a face mask is one of the main Preventive measures among them. Wearing a Mask too frequently can reduce the risk of COVID-19 transmission. However, it is difficult to expect that everyone is able and prepared to wear a mask. To overcome this above challenge Many Artificial Intelligence Researchers come with a solution to classify people with masks and without a mask. This helped Government and Health Workers to Identify the Violated Zones in their area and also to find out people who came in contact with covid19 Positive patients. This helped Lot to avoid the spread of Covid19 patients.

II. DESCRIPTION OF THE PROBLEM

2.1 Statement of Problem

Public places are quite filled with crowded people sometimes. The increased crowd consists of Lots of faces in Video Surveillance. This makes things complex to identify each and every face in the video. Increased faces make things delayed. Which might Increase the spread of Coronavirus fast.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

In 2020 JingxiaoZheng and his team came with a new Research Paper at IEEE entitled "An Automatic System for Unconstrained Video-Based Face Recognition" [1]. The Paper has research on real-time face recognition in video streams. This helps in identifying recording conditions.

Considering the above-mentioned problems in mind we decided to design a system that will help to classify the people with masks and without a mask and also will recognize and notify the people who are not wearing masks. This will reduce the spread of coronavirus in public places which are under CCTV surveillance.

III. LITERATURE REVIEW

Haar cascade frontal face is an Object Detection Algorithm used to identify faces in an image or a real-time video [2]. The Haar cascade Object Detection Algorithm uses line or edge identification features presented by Viola and Jones in their research paper "Rapid Object Detection using a Boosted Cascade of Simple Features" published in 2001 [2].

YOLO stands for You Only Look Once. YOLO is an effectual real-time object recognition algorithm. YOLO was first Presented in the seminal 2015 paper by Joseph Redmon [3]. YOLO work on neural network for giving us real-time object detection.

Among these, both YOLO algorithm is in demand because of their accuracy and speed. YOLO is used in several fields to detect Buses, people, traffic signs, speed limit signs, and animals. So, we decided to move with YOLO for face detection.

After Detecting the face, we use CNN to classify the face with facemask and Without facemask. We studied 3 Pretrained CNN namely VGG-19, Inceptionv3 (GoogLeNet), resnet50. VGG-19 is a convolutional neural network. VGG-19 has 19 layers of convolutional neural network. VGG-19 has 1 SoftMax layer, 3 fully connected layers, 5 MaxPool layers, and, 16 convolution layers. Inceptionv3 was a CNN with 48 deep layers. ResNet50 is a form of the ResNet model. ResNeet50 is 48 Convolution layers deep. ResNet50 has 1 Max-Pool and 1 Average Pool layer. ResNet50 includes 3.8 ∗109 Floating points operations. To Reduce Computational Power for the Proposed System we will be using deep convolutional neural networks based on VGG-19 pre-trained for object detection tasks on the ImageNet dataset.

IV. SYSTEM DESIGN AND FLOW

In this Project, Python Framework and Machine Learning Will be used. Python will be used in frontend as well as in backend. Python Will also be used for machine learning task and which will help in using TensorFlow library for machine learning. TensorFlow will help in building CNN architecture by using keras module inside it. First of All, we will take video from CCTV. Each frame of video will be taken. Processing of Frame will be done. Later on, the frame will be shared with yolo algorithm for face detection. Once we get the co-ordinate of face, we will detect masks and classify them in two classes. The classes are with mask and without mask. The faces with masks will be skipped then. And faces without mask will go under face recognition phase. Face Recognition will use Google FaceNet Algorithm for Identifying face which are stored already inside system database. After recognizing face, the email notification will be sent to that person whose face is identified. Which will notify him for wearing mask or action against him will be taken.

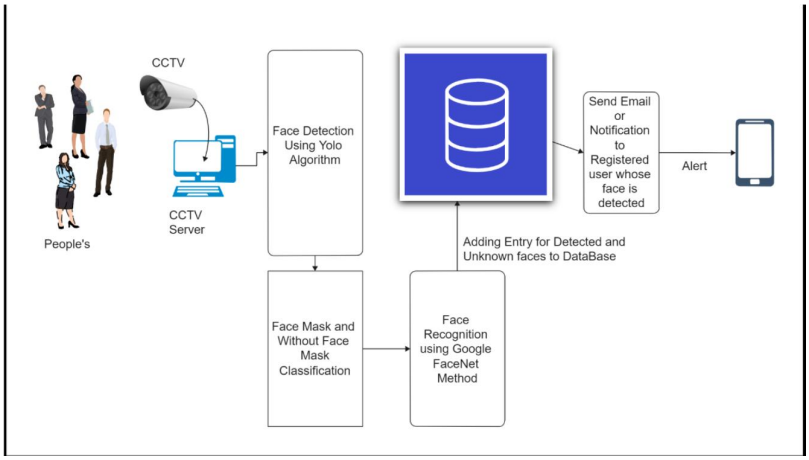

Fig. System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

You only Look Once (YOLO)

You only look once (YOLO) is a modern, real-time object identification program based on Deep learning. Its authors describe how itworks: Prior detection systems repurpose classifiers or localizers to perform detection [7]. They use the model in the image in many areas and scales. Highlights of high scoring inside image are considered detection of object [7].

YOLO researchers use a completely different approach. They use a single neural network in the full image. This network divides the image into regions and predicts bounding boxes and probability for each region. These bounding boxes are weighted with predicted Probability [7].

YOLO has a few advantages over category-based systems. It looks at the whole picture during testing so that its predictions are informed by the global context in the image. It also makes predictions on a single network test unlike systems like R-CNN that require thousands of a single image. This makes it much faster, 1000x faster than R-CNN and 100x faster than Fast R-CNN [7].

The original Yolo model can identify up to 80 different classes of objects with high accuracy. We used this model to find only one thing - the face. We will be training YOLO on WiderFace dataset (a dataset containing images with 393,703 face labels). There is also a smaller version of the Yolo algorithm available, Yolo-Tiny. Yolo-Tiny takes less time to calculate by compromising its accuracy. Engati.com says they trained the Yolo-Tiny model with the same database, but the bounding box results were inconsistent [8].

Convolutional neural network (CNN)

With Face mask and without face mask classification is performed using CNN's binary image classification method to check the presence of face mask on detected faces. A number of models have been developed using CNN to differentiate With Face mask and without face mask into 224*224 images and their performance in terms of recall, F1 scores, accuracy, and precision compared to class 0 (no mask) and class 1 (masked). VGG19 was selected because of its performance during prediction and accuracy. A pre-trained imagnet model is used in conjunction with VGG19.

Google FaceNet

FaceNet is the face recognition system. FaceNet was developed by Google Researchers in 2015 [3]. They proposed how to produce a high-quality facial map from photographs using in-depth reading formats such as ZF-Net and Inception. He then used a method called triplet loss as a function of loss to train the Architecture. Let's look at the architecture in more detail.

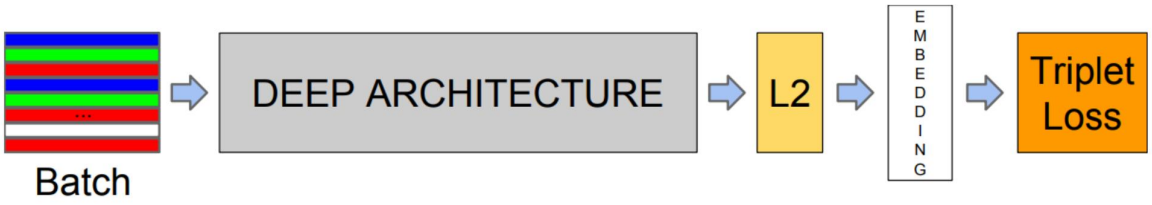

Fig. FaceNet Architecture

FaceNet uses end-to-end learning in its construction. It uses ZF-Net or Inception as its basic architecture. Also add a few 1 * 1 convolutions to reduce the number of parameters. These deep learning models outputs an embedding of the image f(x) with L2 normalization performed on it [3]. This is embedded and transferred to the loss function to calculate the loss. The goal of this loss function is to make the square distance between the two images embedded independent of the image state and the shape of the same identity is smaller, while the square distance between the two images of different identities is greater. Hence, Triplet loss is used as a new function. The idea of using a triple loss in our design is that it makes the model force the boundary between the faces of the various identities [3].

The face-to-face combination (photo collection of the same person) from the FaceNet page shows that the model does not change in size, shape, brightness and even age, etc [3].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

VI. PROJECT IMPLEMENTATION

Face mask Detection and Face recognition are Individual Systems. Considering this in mind weunderstood the gap of technology.Face mask Detection uses a Convolutional neural network. In deep learning, a convolutional neuralnetwork is a class of artificial neural network, most commonly applied to analyze visual imagery[13]. It is based on the shared-weight architecture of the convolution kernels or filters that slidealong input features and provide translation equivariant responses known as feature maps [13].

Before Detecting a face mask the face detection is an important task. The previous systemmainly used the Haar cascade object detection algorithm for face Detection. This was a bit timeconsuming. So we decided to come with YOLO Algorithm which is a faster and real-time object detection algorithm. The YOLO uses Neural Network for Detecting objects in video frames as wellas in Images. YOLO has a very good accuracy than any other object detection algorithm now. It is most recent Algorithm used for Object detection and has came up with various version of itsown.

After classifying faces with masks and without masks the next task is to recognize the faceswithout masks. This system is going to use the most efficient algorithm for face recognition namedas FaceNet. FaceNet is the name of the facial recognition system that was proposed by Google Researchers in 2015 in the paper titled FaceNet: A Unified Embedding for Face Recognition andClustering. It achieved state-of-the-art results in the many benchmark face recognition dataset such as Labeled Faces in the Wild (LFW) and Youtube Face Database.

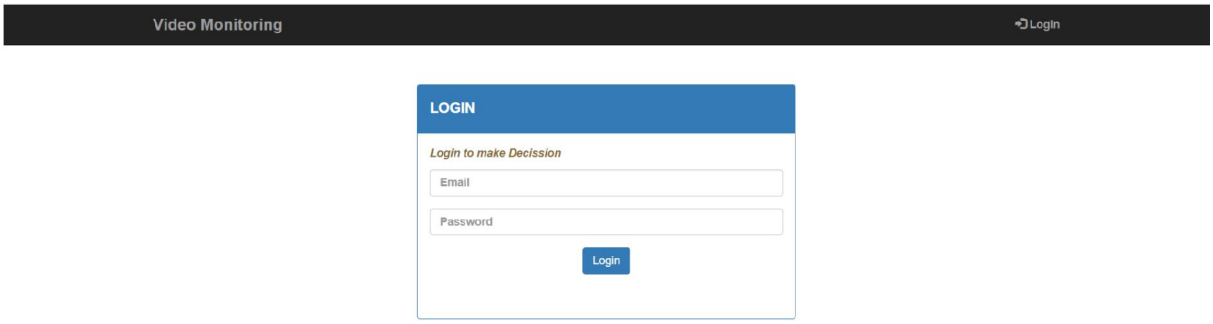

Fig. Admin Login GUI

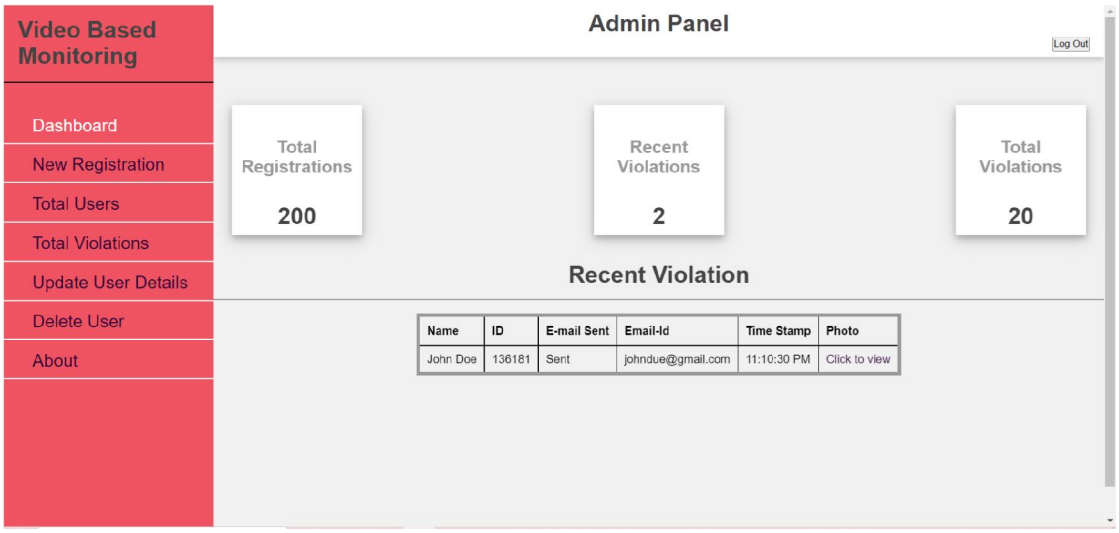

Fig. Dashboard GUI

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

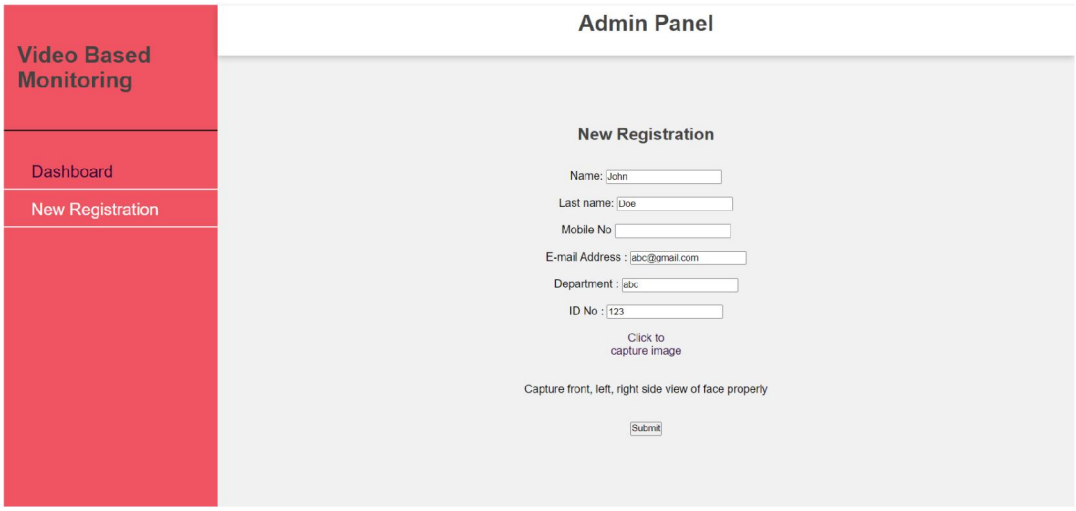

Fig. New Person Registration GUI

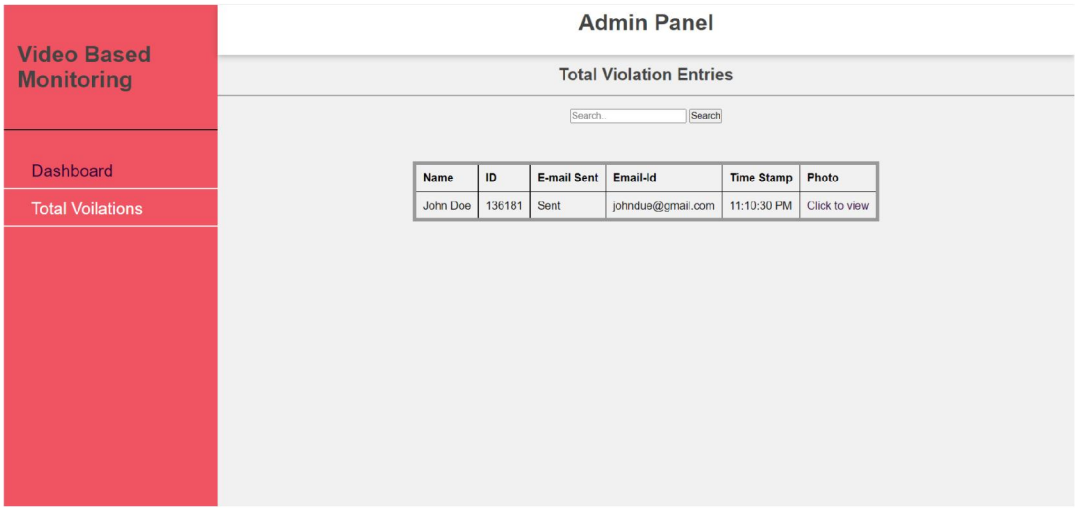

Fig. Total Violation GUI

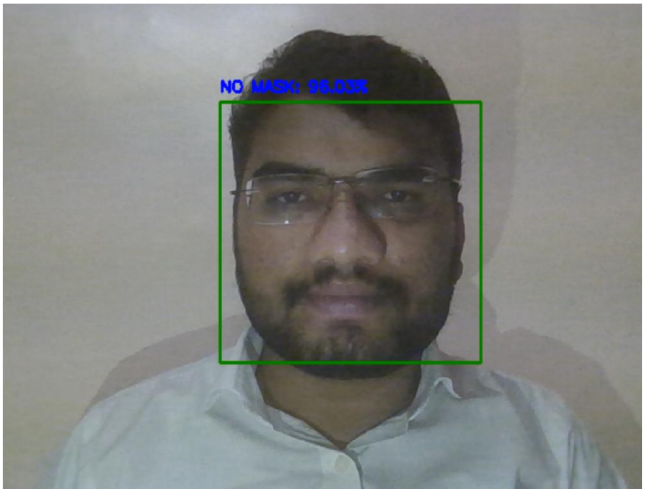

Fig. Video Stream GUI without mask Prediction

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

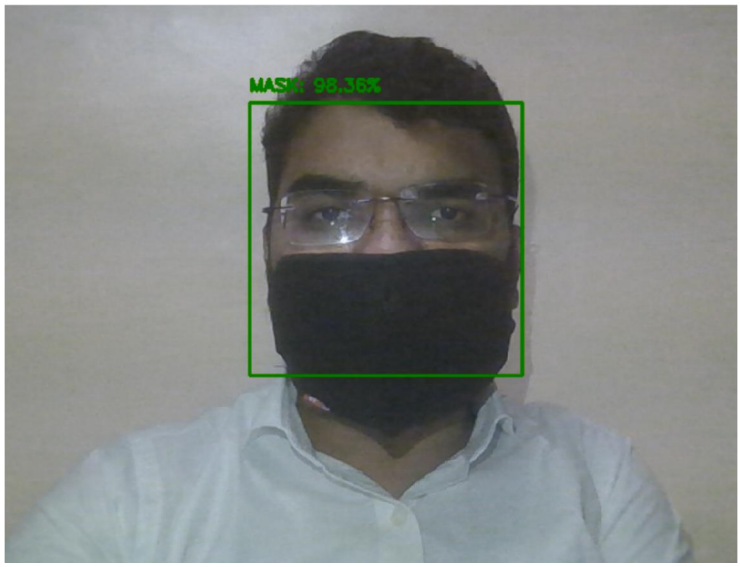

Fig. Video Stream GUI with Mask Prediction

VII. DISCUSSION AND RESULTS

Training Machine Learning Model

For the Classification of people's Faces with a mask and without a mask, we will need to trainthe Machine Learning Model Using CNN Algorithm. Also For Face Recognition, we will need totrain the Machine Learning model by using Google FaceNet Algorithm. Designing Neural Networkfor training and getting higher accuracy in this system takes a long time.

Implementing Functions of System

After Designing the Database, GUI, and Machine Learning model next step is to merge all thesethree and Create a Complete System by using various API. The Role Of the BackEnd Engineercomes in this task. one can also Deploy this project on the cloud and create a simple CI/CDpipeline for the tasks which occur recursively.

Testing

After Implementation of the whole System, One needs to Test the System Under various Conditions.The Tester Does this task. The Outcome of testing Should be finding bugs in the Systemif Any Exist. If the System is Bug-Free and Ready to Use it will be deployed to the Client.

Reporting

Every System should have Documentation along with it. The Report is Created as Documentationof the System. The Report of the Project contains all the details of the System from theRequirement Gathering to the Deployment of System. It is simply a guide to a person who wantsto design another similar system.

VIII. ADVANTAGES & DISADVANTAGES

Advantages:

- Very accurate, flawless. Sooner than MTCNN [8].
- Intelligent Alert:This System Alerts People on their smartphones by using email notifications. This is Done bythe Classification of People's faces with mask and without a mask and after that identifying thatfaces. This operation is done without any human interaction with the system, Hence this is an example of the self intelligence system and so we call it an Intelligent alert system.
- Camera Agnostic:The Main Input Device of this system is the Camera. The System Supports any type of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

Camera which is connected to your computer or server. The Camera outputs images or video and thisoutput is provided to the system by the camera. And System mainly uses videos and images forfurther operations hence this system is indirectly a Camera Agnostic.

- Maximum Accuracy: Accuracy of a System is a percentage of a system giving true results. Maximum accuracy givesthe probability that the result is right. This System Gives around 96% Accuracy for FaceMask Detection using CNN Model.
- Improved security: Face Recognition helps in increasing the effort of a Surveillance system. This System Keeps the Records of Unknown People who are not using the mask which helps in increasing Security in theProvince.

Disadvantages:

- As it has large Deep Neural Network layers, it needs additional computational resources. Therefore, slow operation on CPU or mobile devices is slow. On the GPU, it takes up more VRAM due to its larger architecture.
- Privacy: There have been a few instances in the past where security cameras have stirred up controversies, especially in professional setups [4]. There have been cases where employees have objected tobeing under constant surveillance without their permission and citing the ‘invasion of privacy asthe reason [4]. A few have also resorted to taking legal action against their employers concerning this [4].
- Massive Storage Capacity: Face Recognition Stores a Large number of images. To Store these Images the System needshigh storage capacity. This Storage may not be available to every organization.

Applications

- Security and Defense: The System Detects unknown persons in the organization which will help in Security and Defensefor recognizing such unknown people and avoids threats for normal People from such unknownPersons.
- Face Recognition for Access Control: The Face Recognition Algorithm used in the system can help the organization to give Access toPeoples at Different Restricted areas in an organization.
- Surveillance system: The CCTV Surveillance System keeps every record of action Performed under CCTV camerawhich helps to detect a threat in the organization and help in taking action against such threats.

IX. CONCLUSION & FUTURE WORK

There is a dearth of datasets to be used for such a system and it is not diverse to work for allthe situations [11]. The social distancing calculation and Rule Violation can also be implementedin the same system. For instance, the system sometimes confuses beards with masks due to nothaving enough negative examples with beards in it [11]. When such datasets become available, amore powerful model can be trained [11]. The Increase in Frames per second than 30fps may lead to a frame drop in the system so there is the chance to design a system to overcome this issue.

This report provides an efficient solution to monitor mask detection practices in public areaswhere it is very difficult to monitor manually. Four different modules have been developed for facedetection, face mask classification, face Recognition, and Email Notification. The system performs reasonably well with an accuracy of 96% for face mask classification. Here face detection takesplace using YOLO v4. Without any addition of time-consuming computations or image warping,this lightweight model is easy to calibrate and can be well used in real-time due to 30FPS andgood accuracy. The Project results demonstrate the superior performance of the proposed system

ACKNOWLEDGEMENTS

The completion of our project brings with it a sense of satisfaction, but it is never complete without those people who made it possible and whose constant support has crowned our efforts with success. One cannot even imagine our completion of the project without guidance and neither can we succeed without acknowledging it. It is a great pleasure that we acknowledge the enormous assistance and excellent co-operation to us by the respected personalities.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, May 2022
Impact Factor: 6.252

REFERENCES

[1]. J. Zheng, R. Ranjan, C.-H. Chen, J.-C. Chen, C. D. Castillo and R. Chellappa, "An automatic system for unconstrained video-based face recognition," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 2, pp. 194--209, 2020.
[2]. G. S. Behera, "Face detection with haar cascade," December 2020. [Online]. Available: https://bit.ly/3E2mt6p.
[3]. GeeksforGeeks, "Facenet – using facial recognition system," [Online]. Available: https://www.geeksforgeeks.org/facenetusing-facial-recognition-system/.
[4]. K. Baluprithviraj, K. Bharathi, S. Chendhuran and P. Lokeshwaran, "Artificial Intelligence based Smart Door with Face Mask Detection," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 543--548, 2021.
[5]. S. Srinivasan, R. R. Biradar, S. Revathi and R. R. Singh, "COVID-19 Monitoring System using Social Distancing and Face Mask Detection on Surveillance video datasets," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 449--455, 2021.
[6]. K. Bhambani, T. Jain and K. A. Sultanpure, "Real-time Face Mask and social distancing violation detection system using YOLO," 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), pp. 1--6, 2020.
[7]. J. Redmon and A. Farhadi , "Yolov3: An incremental improvement," ar Xiv preprint arXiv:1804.02767, 2018.
[8]. Engati, "How will facial recognition systems algorithms work in 2022," 2021. [Online]. Available: https://www.engati.com/blog/facial-recognition-systems.