Grenze International Journal of Engineering and Technology, Jan Issue

Face Mask and Social Distance Detection at Work Place

G K Sreevineeth[1], Shivaneeth P[2] and Kavitha M[3]
[1-3]Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India
Email vsreevineeth@gmail.com, shivu61133@gmail.com, kavitham@sit.ac.in

***Abstract*—Coronavirus, the cause of COVID-19 epidemic, is causing a global health catastrophe. Wearing a facemask and maintaining physical distance are effective protection measures against COVID-19. World health organization suggests that, while in public places everyone should keep at least one meter away from each other. Wearing facemasks at work place significantly reduces the risk of virus transmission. This scenario forces the world community to seek alternative strategies of limiting the spread of this infectious illness. Social distance is the one of the viable strategies for combating this epidemic. A face mask is regarded as an important measure of public safety to slow down the transmission. In the current context, it is vital to foster this trait. With this motive of monitoring COVID-19 Safety guidelines at work place, we have developed a tool that identifies face masks and monitors social distancing using deep learning techniques. In the current scenario, there is a huge necessity to build these types of tools to monitor people at work place. The combination of Artificial Intelligence and deep learning has previously been used to produce promising outcomes in solving a variety of everyday problems. This tool analyses live video streams from cameras and determines whether people are keeping a safe distance from one another. Simultaneously, the face mask will be identified in the video frame. Security cameras at universities and other workplaces can be integrated with this tool to keep a careful eye on people's security measures and take appropriate action when required.**

***Index Terms*— Coronavirus, epidemic, facemask, social distance, workplace, live video streams, computer vision, deep learning.**

I. INTRODUCTION

We are currently in a global pandemic situation caused by COVID – 19, also known as novel corona virus. It is linked to the SARS (severe acute respiratory syndrome) respiratory virus. According to the World Health Organization, the virus was discovered on December 31st, 2019 in the city of Wuhan, Hubei China, with unknown origin or source. In a few of months, the virus had spread all over the world. To prevent the spread of the virus, every country proclaimed a state of emergency. Because it is a respiratory ailment, it can easily spread through the air. In normal circumstances, the virus can survive in the open air for one to two days. Cough, fever, headache, odour loss, weariness and exhaustion are common symptoms.
To prevent this, the WHO has announced a few procedures that must be implemented in order to reduce the rate at which the virus spreads. These guidelines include:

- Hands must be watched frequently
- Cleaning with a disinfectant
- Touching the face, eyes, and nose should be avoided as much as possible

Grenze ID: 01.GIJET.9.1.560

- Maintain a minimum physical contact with others
- Mandatory to wear a mask whenever stepped outside
- Keeping a social distance between you and your neighbours of one to two meters in public places

The government has implemented numerous laws and regulations, such as lockdowns, curfews, closing down undesired work forces, and providing work-from-home programmes. The global economy has collapsed, resulting in a slew of economic and financial crises. After about a year of lockdown, the nation's now opened just critical services and allowed organizations to work with a smaller workforce.

For an organization such as, factories or educational institutions to function, requires a large workforce, for which the organisation must take safeguards to ensure the workers' safety. According to study, wearing a mask at work greatly minimises the chance of viral transmission. Distancing oneself from the others has proven to be an extraordinarily efficient way of limiting its progress.

When a facemask is worn correctly, regardless of whether the individual has symptoms or is asymptomatic, the spread of viruses to another can be prevented.

Many countries came out of lockdown once the first wave ended. Despite restrictions and norms, many people overlook the significance of maintaining a social distance and wearing a face mask. The propagation of the second wave, on the other hand, began with the onset of negligence. Many people will follow the laws only if they are being watched, yet watching everyone in every place at all times would be both time intensive and impossible.

We created a monitoring mechanism to help us overcome this problem. Social distance monitoring and face mask detection have been implemented in our project. Live footage can be collected through the organization's CCTV or any other camera connected to it. Many firms already have CCTV systems in place, making our product simple to implement. If our system detects someone who is not adhering to the rules of covid-19 inside the work place, an automatic warning message will be delivered to that individual.

Our system is suitable for use in a variety of work environments, including offices, educational institutions, and other organisations. This allows us to eliminate physical contact among co-workers in the organisation with minimal effort and expense.

Existing solutions-

- At the entrance or exit points of any organisation or public location, a human force such as a security guard or police officer is assigned the job of checking the temperature and a facemask.
- Facemask violators are fined by the relevant authorities.

Limitations of existing solutions-

- Temperature and facemask monitoring is done exclusively at entry and exit locations.
- Within the organization, there is no continual monitoring surveillance of facemasks or social distance
- It is imperative that humans intervene
- Any human error, such as his absence or just neglecting his responsibilities, could occur.

Proposed solution-

- Face mask detection and social distance monitoring using an automated approach
- The proposed approach can be implemented on existing CCTV cameras in an enterprise at a reasonable maintenance cost.
- It is possible to identify the specific person who is not adhering to the rules of covid-19 and take the appropriate measures.
- Human intervention is kept to an absolute minimum.
- The proposed method outperforms the present system in terms of efficiency.

II. Proposed Algorithm

The six phases of this monitoring process are outlined below.

A. Obtain video footage

This model can be implemented for both real-time video recording and pre-recorded video.

Live video can be captured with a CC camera or IP camera of a specific organization.

In order to access mobile video footage, an application "Ip webcam" must be preinstalled on android mobile. An Ip address must be supplied as input to the system. Both the mobile phone and the monitor must be connected to the same network to transmit data.

Figure 1: Control flow of the application

B. Person detection

This is the first and most important phase in the process of monitoring.
Initially, an object detection algorithm is used to detect all individuals (and only people) in a video stream. The (x, y) coordinates of a person in the frame are determined by SSD using Mobile net model.

Object detection

The term "object detection" refers to a computer vision technique that identifies and locates objects in images and videos. Object detection deals with the classification and localization of objects. Image classification and object localization are two different tasks, which are combined to achieve the unique goal of object detection.
Image categorization is the process of assigning tags to images based on their intended category. Object localization is the process of recognizing the location of objects in an image and creating an abstract boundary around them.
The object detection technique combines these two tasks by drawing a bounding box around each object and assigning it a class label to detect and classify the object in the image. This entire process is known as object recognition.
The task of locating all human instances in an image is known as human detection. It entails looking for all spots in the image at all scales and matching the little region of each position to a template or pattern that the person is familiar with. Human tracking is the technique of generating a persistent path or trajectory for a person by linking human detections in a video series across time. Human detection and tracking are the first two steps in the video surveillance pipeline, and they can be sent into higher-level inference modules.
Algorithm-1: Mobilenet-ssd-
SSD (Single Shot MultiBox Detector) is a popular object detection algorithm and the mobile network is a convolutional neural network with advanced features. The Mobilenet-ssd model is a Single Shot MultiBox Inspection (SSD) network used for object detection. This model is implemented using the Caffe * framework. By using SSD, we can detect multiple objects in the image with a single shot. The architecture of the model is based on the inverted residual structure. Unlike the traditional residual model, the entry and exit of the residual block are thin bottleneck layers. In addition, in the interlayer, the nonlinearity is removed and a slight deep convolution is applied. This model is part of the TensorFlow Object detection API.
Single Shot Detector (SSD) is based on a set of predetermined areas. A grid of anchor points is placed on the input image. At each anchor point, the boxes of various shapes and sizes are used as areas. For each box at each anchor point, the model generates a prediction of whether there is an object in the area and modifies the position and size of the box to bring it closer to the object. Because there are multiple boxes for each anchor point and the anchor points can be very close together, the SSD will generate many potential overlapping detections. Post-processing should be applied to the SSD output to select the best prediction.

C. Social distance detection and monitoring

In addition to the (x, y) coordinates of a person in the frame, which was obtained from object detection algorithm, the distance from the camera to the person is considered as Z coordinate such that 3-dimentional coordinates are obtained. After centroid calculation, Euclidean distance formula is applied to get approximate pixel distance among the persons recognized in the frame.
Algorithm-2: Triangle similarity technique-
The first step in finding the distance to an object in the image is to calculate the focal length. For this, we need to know:

- The distance from the camera to the object.
- The width of this object (in inches, meters, etc.).

Calculation of the focal length of the camera
To determine the distance from our camera to a known object, we have used the similarity of triangles.
suppose we have an object of width W. Then we place the marker at a certain distance D from our camera. We use the camera to take a picture of the object and then measure the apparent width in pixels P. This allows us to deduce the perceived focal length (F) of the camera:
F = (P x D) / W
Form the known P, D and W, we can calculate F, then use it and find the distance D` of an object with a width of P pixels.
D` = (width x depth) / P
Improving distance measurement with camera calibration-

To perform distance measurement, we must first calibrate our system. In this project, we have used a simple "pixels per measurement" technique. However, performing proper camera calibration by calculating external and internal parameters can achieve higher accuracy.

• External parameters such as translation and rotation of matrices are used to convert something from the world frame to the camera frame

• Internal parameters are camera parameters, such as focal length, convert this information to pixels

The most common method is to use OpenCV to perform a checkerboard camera calibration. Doing so will eliminate radial distortion and tangential distortion, both of which will affect the output image, thereby affecting the output measurement of the object in the image.

D. Face and facemask detection

After successful human detection, face of a person is identified.

A hybrid model using deep and classical machine learning for face mask detection is integrated with this tool. A face mask detection dataset consists of, with mask and without mask images, we have used OpenCV to do real-time face mask detection from a live stream via an Ip camera

Algorithm – 3: DNN Face Detector in OpenCV-

One of the most fundamental parts of computer vision is face detection. Many follow-up studies, such as finding individual persons to mark crucial points on the face, are based on it. An artificial neural network (ANN) having numerous layers between the input and output layers is known as a deep neural network (DNN). It's a Caffe model that uses the ResNet10 architecture as its backbone and is built on Single Shot Multiple Box Detector (SSD). After OpenCV 3.3, it was added to its deep neural network module. Although a quantized version of TensorFlow is available, we chose Caffe's model for this project.

The components of DNN face detector are:

neurons, synapses, weights, biases, and functions.

These components work in a similar way to the human brain and can be trained like any other ML algorithm.

For example, a DNN trained to recognize dog breeds will examine a given image and calculate the probability that a dog in the image belongs to a specific breed. The user can view the results and select the probabilities to be displayed on the network (above a certain threshold, etc.), as well as return a suggested label. Each math operation is considered a layer in itself, and a complex DNN has many layers, therefore it is called a "deep" network.

DNN networks are typically a feedback network where data is passed from the input layer to the output layer without Return. First, DNN creates a virtual graph of neurons and assigns random values or "weights" to the connections between them. The weight is multiplied by the inputs and returns a result between 0 and 1. If the network cannot accurately identify the specified pattern, the algorithm adjusts the weights. In this way, the algorithm can increase the influence of certain parameters until it determines the correct math operation for complete data processing.

E. Identification of Violators

- Each frame from a particular video footage is dynamically updated with the changing colors of the boundary boxes, around the person and his face.
- The red bounding box around a person's face indicates that he is not wearing the face mask and is also moving in close proximity to others, breaking the social distance norm.
- The person is being safe by following all covid-19 guidelines, as indicated by the green boundary box.
- The color variation in the box helps the CCTV operator to determine the frequency of violations in a specific area and can alert the relevant authorities.

Algorithm – 4: Face net model-

The process of recognizing and authenticating people through photographs of their faces is known as facial recognition. Face Net is a facial recognition system developed by Google researchers Florian Schroff, Dmitry Kalenichenko, and James Philbin in the article Face Net: A Unified Embedding for Face Recognition and Clustering, which was published in 2015.

One of the most intriguing uses of deep learning is facial recognition. Face Net unifies facial recognition, verification, and grouping operations into a unified architecture. It achieves state-of-the-art precision by employing deep convolutional networks and triple loss. Each image of a face is mapped in Euclidean space so that its distance in space correlates to its resemblance to the image of any other individual.

The first step in creating a face recognition system is establishing the bounding box of face position. Then, find the spatial points of facial characteristics like the length of the eyes, the length of the mouth, the distance

between the eyes and the nose, and so on. Embeddings were created using retrieved face characteristics and the distance between two embeddings would be determined.
Face Net does not define any new algorithms to perform the above tasks, but only creates embeddings, which can be used directly for face recognition, verification, and grouping. Face Net uses a deep convolutional neural network (CNN). The network is trained so that the distance L2 squared between inlays corresponds to the similarity of faces. Images used for training are precisely scaled, transformed and cropped around the face area.

Algorithm – 5: Support vector machines-

Support Vector Machine (SVM), is a machine learning algorithm that can be used for both regression and classification. Many people prefer support vector machines because they generate high precision with little computing power

The support vector machine technique seeks a hyperplane in an N-dimensional space (N = number of features) that delimits data points unambiguously. The hyperplane is a decision boundary that aids in data point classification. Data points on either side of the hyperplane are almost certainly assigned to separate classes. There are numerous hyperplanes from which to split the data pieces of the two classes. Our goal is to locate the plane with the greatest margin, or the greatest distance between two types of data points. Maximizing the margin distance gives some reinforcement for classifying subsequent data points more consistently.

F. E-mail Alert

An email alert will be sent to the person who are not wearing facemask and whose face is recognizable in organization database. The database is dynamically updated with a list of violators, together with their images and violation details.

III. Experiment and Result

The experimental setup in which this model has been implemented is described as follows:
Operating System: Windows 10
Disk space 1TB
Intel® CoreTM i5 processor 6200U at 1.60 GHz or 1.80 GHz, 8 GB of DRAM
The proposed system is tested across multiple video footages and the results are shown below.

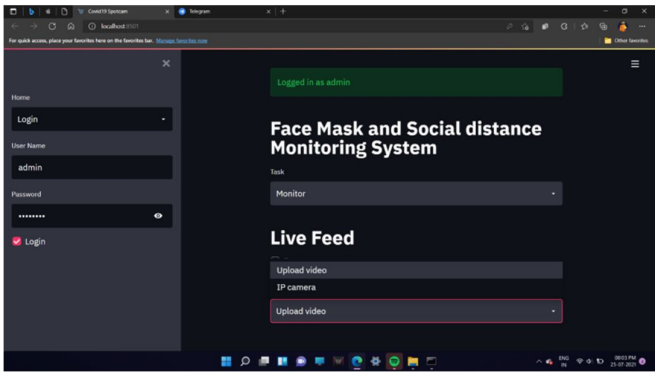

Figure 1: Application home page along with admin portal

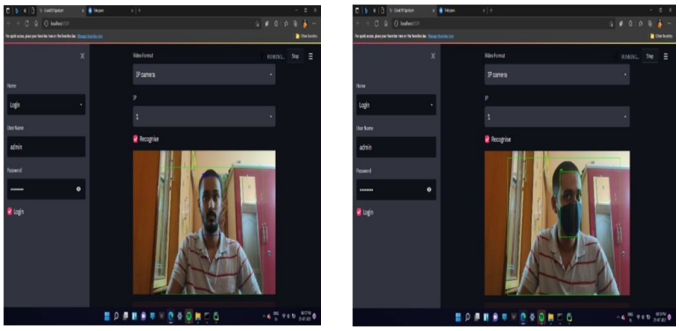

Figure 2: Face mask detection with bounding box

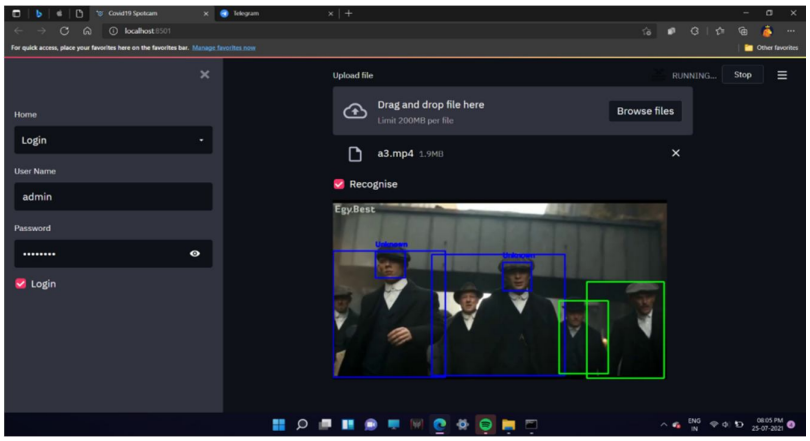

Figure 3: Running tool on local video footage and social distance monitoring

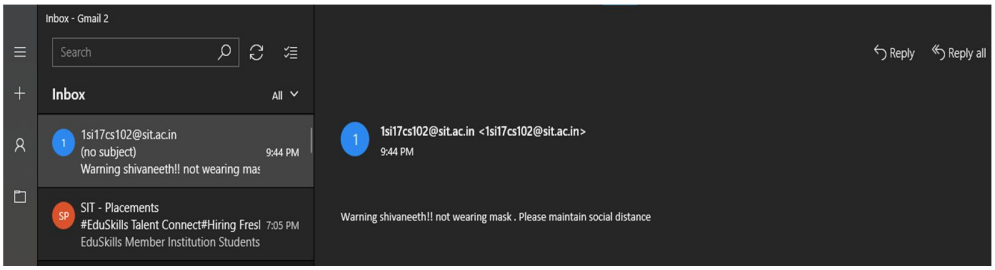

Figure 4: Email alert notification

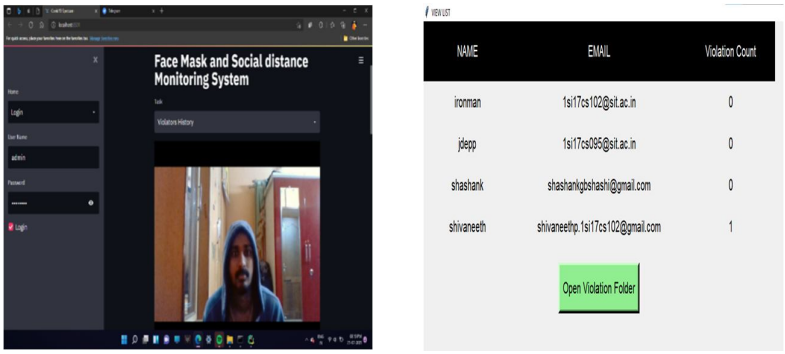

Figure 5: violators history

IV. CONCLUSION

The fundamental goal of this project is to employ artificial intelligence to make workplaces, factories, schools and universities safer places for people. As there is a spike in number of cases on a daily basis, this type of AI-based solution becomes increasingly necessary. This tool can be integrated with a company's security camera system, which can track whether, people are keeping a safe distance and predict the level of risk at work place. This method is more useful for close monitoring of people in covid-affected micro containment regions. When used properly, this monitoring system can quickly alert appropriate authorities. As a result, people will adhere to covid safety protocols at all times.

REFERENCES

[1] "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and deep sort techniques" published at arXiv:2005.01385v2 [cs.CV] 6 May 2020
[2] "Retina face mask: A face mask detector" published at arXiv:2005.03950v2 [cs.CV] 8 Jun 2020

[3] “A vision based social distancing and critical density detection system for COVID-19” published at Group Data & Analytics, ABMCPL, Aditya Birla Group, India
[4] “Deep learning based safe social distancing and face mask detection in public areas for COVID19 safety guidelines adherence” published at International Journal for Research in Applied Science & Engineering Technology (IJRASET)
[5] “Facial mask detection using semantic segmentation” published at 2019 4th International Conference on Computing, Communications and Security (ICCCS)