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

COVID-19-Preventions-Control-System, Face-Mask, And Face-Hand Detection Framework

Dr. S. Sathyalakshmi[*1],Medikonda Venkata Leela Krishna[*2], Phani Vineel Varma[*3], Inkulu Sujith[*4], Pelluri Venkata Sai Goutham Aditya[*5]
[*1]Associate Professor, [*2,*3,*4&*5] B. Tech Student
[*1,*2,*3,*4&*5]Hindustan Institute of Technology and Science, Kelambakkam, Tamil Nadu , India

Article Info

Publication Issue :
Volume 8, Issue 6
November-December-2022

Page Number : 147-155

Article History
Accepted: 05 Nov 2022
Published: 15 Nov 2022

ABSTRACT

The end of 2019 witnessed the outbreak of Coronavirus Disease 2019 (COVID-19), which has continued to be the cause of plight for millions of lives and businesses even in 2021. As the world recovers from the pandemic and plans to return to a state of normalcy, there is a wave of anxiety among all individuals, especially those who intend to resume in person activity. Studies have proved that wearing a face mask significantly reduces the risk of viral transmission as well as provides a sense of protection. However, it is not feasible to manually track the implementation of this policy. Technology holds the key here. So we are introducing a system based on deep learning that which can identify the person either wearing a mask properly or not. To implement the process we consider the dataset called MAFA-data which will be trained using Convolution neural network (CNN) along with computer vision.

Keywords : Deep Learning, Convolutional Neural Network, Face mask image dataset

I. INTRODUCTION

Rapid advancements in the fields of Science and Technology have led us to a stage where we are capable of achieving feats that seemed improbable a few decades ago. Technologies in fields like Machine Learning and Artificial Intelligence have made our lives easier and provide solutions to several complex problems in various areas. Modern Computer Vision algorithms are approaching human-level performance in visual perception tasks. From image classification to video analytics, Computer Vision has proven to be a revolutionary aspect of modern technology. In a world battling against the Novel Coronavirus Disease (COVID-19) pandemic, technology has been a lifesaver. With the aid of technology, 'work from home' has substituted our normal work routines and has become a part of our daily lives. However, for some sectors, it is impossible to adapt to this new norm. As the pandemic slowly settles and such sectors become eager to resume in-person work, individuals are still skeptical of getting back to the office. 65% of employees are now anxious about returning to the office (Woods, 2020).

Multiple studies have shown that the use of face masks reduces the risk of viral transmission as well as provides a sense of protection (Howard et al., 2020; Verma et al., 2020). However, it is infeasible to manually enforce such a policy on large premises and track any violations. Computer Vision provides a better alternative to this. Using a combination of image classification, object detection, object tracking, and video analysis, we developed a robust system that can detect the presence and absence of face masks in images as well as videos. In this paper, we propose a two-stage CNN architecture, where the first stage detects human faces, while the second stage uses a lightweight image classifier to classify the faces detected in the first stage as either 'Mask' or 'No Mask' faces and draws bounding boxes around them along with the detected class name. This algorithm was further extended to videos as well.

The detected faces are then tracked between frames using an object tracking algorithm, which makes the detections robust to the noise due to motion blur. This system can then be integrated with an image or video capturing device like a CCTV camera, to track safety violations, promote the use of face masks, and ensure a safe working environment.

2. **Traditional Object Detection**: The problem of detecting multiple masked and unmasked faces in images can be solved by a traditional object detection model. The process of object detection mainly involves localizing the objects in images and classifying them (in case of multiple objects). Traditional algorithms like Hear Cascade (Viola and Jones, 2001) and HOG (Dalai and Trigs, 2005) have proved to be effective for such tasks, but these algorithms are heavily based on Feature Engineering.

A CNN uses convolution kernels to convolve with the original images or feature maps to extract higher-level features, thus resulting in a very powerful tool for Computer Vision tasks. 2.1.3. Modern Object Detection Algorithms: CNN based object detection algorithms can be classified into 2 categories: Multi-Stage Detectors and Single-Stage Detectors. Multi-Stage Detectors In a multi-stage detector, the process of detection is split into multiple steps. A two stage detector like RCNN first estimates and proposes a set of regions of interest using selective search. The CNN feature vectors are then extracted from each region independently. Multiple algorithms based on Regional Proposal Network like Fast RCNN and Faster RCNN have achieved higher accuracy and better results than most single stage detectors.

Single-Stage Detectors A single-stage detector performs detections in one step, directly over a dense sampling of possible locations. These algorithms skip the region proposal stage used in multi-stage detectors and are thus considered to be generally faster, at the cost of some loss of accuracy. One of the most popular single stage algorithms, You Only Look Once (YOLO), was introduced in 2015 and achieved close to real time performance. Single Shot Detector (SSD) (Liu et al., 2016) is another popular algorithm used for object detection, which gives excellent results. RetinaNet (Lin et al., 2017b), one of the best detectors, is based on Feature Pyramid Networks (Lin et al., 2017a), and uses focal loss. 2.1.4. Face Mask Detection: As the world began implementing precautionary measures against the Coronavirus, numerous implementations of Face Mask Detection systems came forth. Have performed facial recognition on masked and unmasked faces using Principal Component Analysis (PCA). However, the recognition accuracy drops to less than 70% when the recognized face is masked. (Qin and Li, 2020) introduced a method to identify face mask wearing conditions.

They divided the facemask wearing conditions into three categories: correct face mask wearing, incorrect face mask wearing, and no face mask wearing. Their system takes an image, detects and crops faces, and then uses SRCNet to perform image super-resolution and classify them. The work by (Nieto-Rodríguez et al., 2015) presented a method that detects the presence or absence of a medical mask. The primary objective of this approach was to trigger an alert only for medical staff who do not wear a surgical mask, by minimizing as many false positive face detections as possible, without missing any medical mask detections. Proposed a model that consists of two components. The first component performs uses

ResNet50 for feature extraction. The next component is a facemask classifier, based on an ensemble of classical Machine Learning algorithms. The authors evaluated their system and estimated that Deep Transfer Learning approaches would achieve better results since the building, comparing, and selecting the best model among a set of classical Machine Learning models is a time-consuming process. 2.2. Proposed Methodology: We propose a two-stage architecture for detecting masked and unmasked faces and localizing them. 2.2.1. Architecture Overview: represents our proposed system architecture.

It consists of two major stages. The first stage of our architecture includes a Face Detector, which localizes multiple faces in images of varying sizes and detects faces even in overlapping scenarios. The detected faces (regions of interest) extracted from this stage are then batched together and passed to the second stage of our architecture, which is a CNN based Face Mask Classifier. The results from the second stage are decoded and the final output is the image with all the faces in the image correctly detected and classified as either masked or unmasked faces.

A face detector acts as the first stage of our system. A raw RGB image is passed as the input to this stage. The face detector extracts and outputs all the faces detected in the image with their bounding box coordinates. The process of detecting faces accurately is very important for our architecture. Training a highly accurate face detector needs a lot of labeled data, time, and compute resources. For these reasons, we selected a pre-trained model trained on a large dataset for easy generalization and stability in detection. Three different pre-trained models were tested for this stage: Dib (Sharma et al., 2016) - The Dib Deep Learning face detector offers significantly better performance than its precursor, the Dib HOG based face detector.

MTCNN (Zhang, K. et al, 2016) - It uses a cascade architecture with three stages of CNN for detecting and localizing faces and facial key points. Retina Face (Deng et al., 2020) - It is a single stage design with pixel-wise localization that uses a multi-task learning strategy to simultaneously predict face box, face score, and facial key points. The detection process is challenging for the model used in this stage, as it needs to detect human faces that could also be covered with masks. We selected Retina Face as our Stage 1 model, based on our experimentation and comparative analysis, covered in section 3.2. 2.2.3. Intermediate Processing Block: This block carries out the processing of the detected faces and batches them together for classification, which is carried out by Stage 2. The detector from Stage 1 outputs the bounding boxes for the faces. Stage 2 requires the entire head of the person to accurately classify the faces as masked or unmasked.

The first step involves expanding the bounding boxes in height and width by 20%, which covers the required Region of Interest (ROI) with minimal overlap with other faces in most situations. The second step involves cropping out the expanded bounding boxes from the image to extract the ROI for each detected face. The extracted faces are resized and normalized as required by Stage 2. Furthermore, all the faces are batched together for batch inference. 2.2.4. Stage 2 - Face Mask Classifier: The second stage of our system is a face mask classifier. This stage takes the processed ROI from the Intermediate Processing Block and classifies it as either Mask or No Mask. A CNN based classifier for this stage was trained, based on three different image classification models: MobileNetV2 (Sandler et al., 2018), DenseNet121 (Huang et al., 2017), Nanette (Zap et al., 2018). These models have a lightweight architecture that offers high performance with low latency, which is suitable for video analysis.

The output of this stage is an image (or video frame) with localized faces, classified as masked or unmasked. 2.2.5. Dataset: The three face mask classifier models were trained on our dataset. The dataset images for masked and unmasked faces were collected from image datasets available in the public domain, along with some data scraped from the Internet. Masked images were obtained from the Real-world Masked Face Recognition Dataset (RMFRD) (Wang, Z. et al., 2020) and Face Mask Detection dataset by Laurel on Cagle (Laurel, 2020).

3.4. Final Results: Combining all the components of our architecture, we thus get a highly accurate and robust Face Mask Detection System. Retina Face was selected as our Face Detector in Stage 1, while the NASNetMobile based model was selected as our Face Mask Classifier in Stage 2. The resultant system exhibits high performance and has the capability to detect face masks in images with multiple faces over a wide range of angles.

3.5. Video Analysis: Until now, we have seen that our system shows high performance over images, overcoming most of the issues commonly faced in object detection in images. For real-world scenarios, it is beneficial to extend such a detection system to work over video feeds as well. Videos have their own set of challenges like motion blur, dynamic focus, transitioning between frames, etc. In order to ensure that the detections remain stable and to avoid jitter between frames, we used the process of Object Tracking. We used a modified version of Centroid Tracking, inspired by (Nescient et al., 1999), in order to track the detected faces between consecutive frames. This makes our detection algorithm robust to the noise and the motion blur in video streams, where the algorithm could fail to detect some objects.

II. RELATED WORKS

Facial Mask Detection using Semantic Segmentation: Face Detection has evolved as a very popular problem in Image processing and Computer Vision. Many new algorithms are being devised using convolutional architectures to make the algorithm as accurate as possible. These convolutional architectures have made it possible to extract even the pixel details. We aim to design a binary face classifier which can detect any face present in the frame irrespective of its alignment. We present a method to generate accurate face segmentation masks from any arbitrary size input image. Beginning from the RGB image of any size, the method uses Predefined Training Weights of VGG – 16 Architecture for feature extraction. Training is performed through Fully Convolutional Networks to semantically segment out the faces present in that image. Gradient Descent is used for training while Binomial Cross Entropy is used as a loss function. Further the output image from the FCN is processed to remove the unwanted noise and avoid the false predictions if any and make bounding box around the faces. Furthermore, proposed model has also shown great results in recognizing non-frontal faces. Along with this it is also able to detect multiple facial masks in a single frame. Experiments were performed on Multi Parsing Human Dataset obtaining mean pixel level accuracy of 93.884 % for the segmented face masks.

Summary: Y. Fang, Y. Nie, and team design a binary face classifier which can detect any face present in the frame irrespective of its alignment.

Face Recognition with Facial Mask Application and Neural Networks: Face recognition represents one of the most interesting modalities of biometric. Due to his low intrusiveness and to the constant decrease in image acquisition cost, it's particularly suitable for a wide number of real time applications. In this paper we propose a very fast image pre-processing by the introduction of a linearly shaded elliptical mask cantered over the faces. Used in association with DCT, for features extraction, and MPL and RBF Neural Networks, for classification, it allows an improvement of system performances without modifying the global computation weight and also a learning time reduction for MLP neural networks.

Summary: S. Feng, C. Shen, N. Xia, and team implement a fast image pre-processing by the introduction of a linearly shaded elliptical mask cantered over the faces.

Face Detection and Segmentation Based on Improved Mask R-CNN: Deep convolutional neural networks have been successfully applied to face detection recently. Despite making remarkable progress, most of the existing detection methods only localize each face using a bounding box, which cannot segment each face from the background image simultaneously. To overcome this drawback, we present a face detection and segmentation method based on improved Mask R-CNN, named G-Mask, which incorporates face detection and segmentation into one framework aiming to obtain more fine-grained information of face. Specifically, in this proposed method, ResNet-101 is utilized to extract features, RPN is used to generate Roils, and Roiling faithfully preserves the exact spatial locations to generate binary mask through Fully Convolution Network (FCN). Furthermore, Generalized Intersection over Union (Giroux) is used as the bounding box loss function to improve the detection accuracy. Compared with Faster R-CNN, Mask R-CNN, and Multitask Cascade CNN, the proposed G-Mask method has achieved promising results on FDDB, AFW, and WIDER FACE benchmarks.

Summary: Z. Wang, G. Wang, B. Huang, and team implement a face detection and segmentation method based on improved Mask R-CNN, named G-Mask, which incorporates face detection and segmentation into one framework aiming to obtain more fine-grained information of face.

Real-Time Face Mask Identification Using Facemask net Deep Learning Network: The COVID - 19 pandemic is devastating mankind irrespective of caste, creed, gender, and religion. Until a vaccine is discovered, we should do our bit to constrain the expanse of the corona-virus. Using a face mask can undoubtedly help in managing the spread of the virus. COVID - 19 face mask detector uses or owns Facemask net, deep learning techniques to successfully test whether a person is with wearing a face mask or not. The manuscript presents three-class classification namely person is wearing a mask, or improperly worn masks or no mask detected. Using our deep learning method called Facemask net, we got an accuracy of 98.6 %. The Facemask net can work with still images and also works with a live video stream. Cases in which the mask is improperly worn are when the nose and mouth are partially covered. Our face mask identifier is least complex in structure and gives quick results and hence can be used in CCTV footage to detect whether a person is wearing a mask perfectly so that he does not pose any danger to others. Mass screening is possible and hence can be used in crowded places like railway stations, bus stops, markets, streets, mall entrances, schools, colleges, etc. By monitoring the placement of the face mask on the face, we can make sure that an individual wears it the right way and helps to curb the scope of the virus.

Summary: Z.-Q. Zhao, P. Zheng, S.-t. Xu, and team focus on facemask net pretrained model to identify the three-class classification namely person is wearing a mask, or improperly worn masks or no mask detected.

A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic: The coronavirus COVID-19 pandemic is causing a global health crisis. One of the effective protection methods is wearing a face mask in public areas according to the World Health Organization (WHO). In this paper, a hybrid model using deep and classical machine learning for face mask detection will be presented. The proposed model consists of two components. The first component is designed for feature extraction using Resnet50. While the second component is designed for the classification process of face masks using decision trees, Support Vector Machine (SVM), and ensemble algorithm. Three face masked datasets have been selected for investigation. The Three datasets are the Real-World Masked Face Dataset (RMFD), the

Simulated Masked Face Dataset (SMFD), and the Labelled Faces in the Wild (LFW). The SVM classifier achieved 99.64% testing accuracy in RMFD. In SMFD, it achieved 99.49%, while in LFW, it achieved 100% testing accuracy.

Summary: A. Kumar, A. Kaur, and team worked on hybrid model using deep and classical machine learning for face mask detection will be presented

III. Methodology

Proposed system:

During pandemic COVID-19, WHO has made wearing masks compulsory to protect against this deadly virus, so our project will notify if someone is wearing mask or not. In this Project we will develop a deep learning. We will use the dataset to build a COVID-19 face mask detector with computer vision using Python, OpenCV, and Tensor Flow and Keras. Our goal is to identify whether the person on image/video stream is wearing a face mask or face hand mask or improper mask or not with the help of computer vision and deep learning.

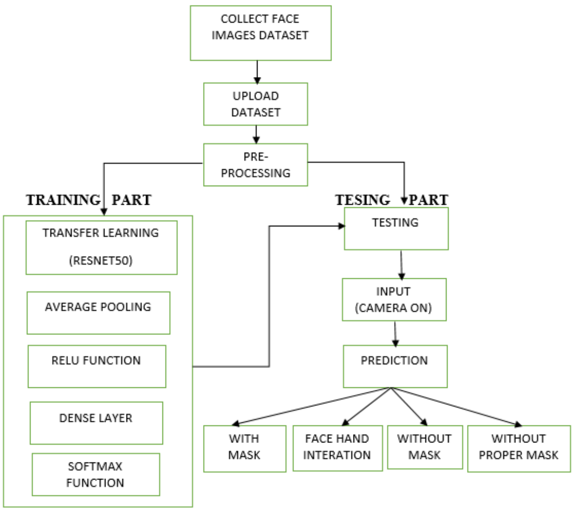


Figure 1: Block diagram of proposed method

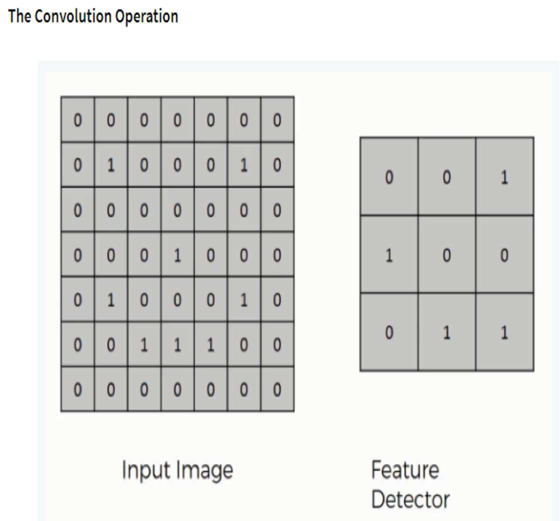
IV. Implementation

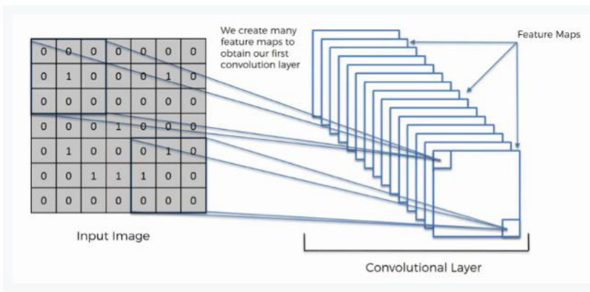
The project has implemented by using below listed algorithms 1.

Convolutional Neural Network

Step1: convolutional operation

The first building block in our plan of attack is convolution operation. In this step, we will touch on feature detectors, which basically serve as the neural network's filters. We will also discuss feature maps, learning the parameters of such maps, how patterns are detected, the layers of detection, and how the findings are mapped out.

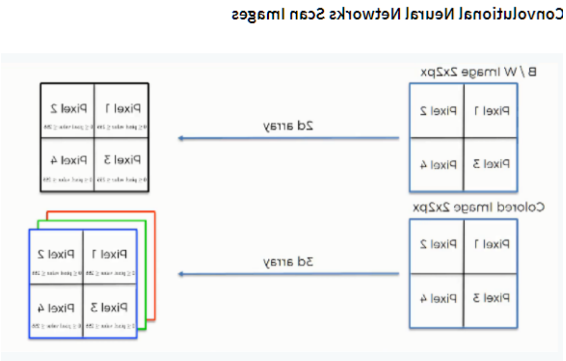




Step (1b): Relu Layer

The second part of this step will involve the Rectified Linear Unit or Relook. We will cover Relook layers and explore how linearity functions in the context of Convolutional Neural Networks.

Not necessary for understanding CNN's, but there's no harm in a quick lesson to improve your skills.



Step 2: Pooling Layer

In this part, we'll cover pooling and will get to understand exactly how it generally works. Our nexus here, however, will be a specific type of pooling; max pooling. We'll cover various approaches, though, including mean (or sum) pooling. This part will end with a demonstration made using a visual interactive tool that will definitely sort the whole concept out for you.

Step 3: Flattening

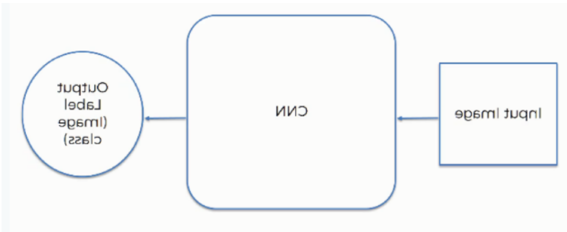
This will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks.

Step 4: Full Connection

In this part, everything that we covered throughout the section will be merged together. By learning this, you'll get to envision a fuller picture of how Convolutional Neural Networks operate and how the "neurons" that are finally produced learn the classification of images.

Summary

In the end, we'll wrap everything up and give a quick recap of the concept covered in the section. If you feel like it will do you any benefit (and it probably will), you should check out the extra tutorial in which Soft ax and Cross-Entropy are covered. It's not mandatory for the course, but you will likely come across these concepts when working with Convolutional Neural Networks and it will do you a lot of good to be familiar with them.

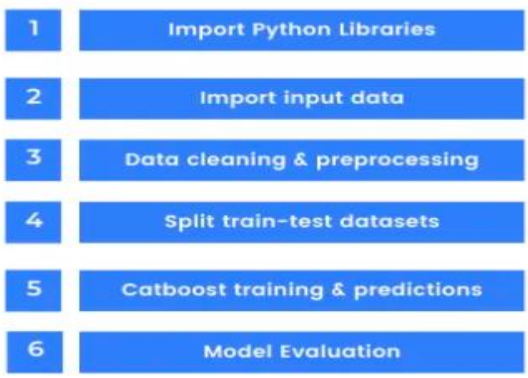


Implementation:

While training time can take up longer than other GBDT implementations, prediction time is 13–16 times faster than the other libraries according to the Yandex benchmark. Catboost's default parameters are a better starting point than in other GBDT algorithms and it is good news for beginners who want a plug and play model to start experience tree ensembles or Kaggle competitions. Some more noteworthy advancements by Catboost are the features interactions, object importance and the snapshot support. In addition to classification and regression, Catboost supports ranking out of the box.

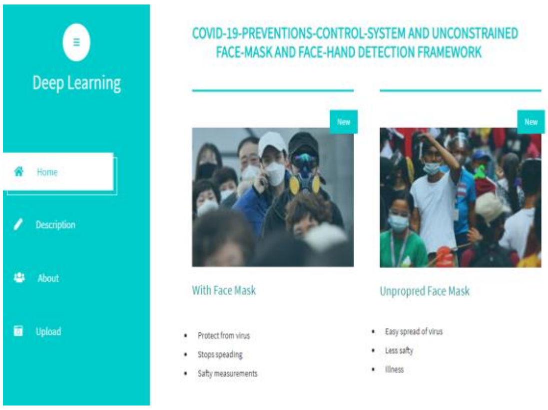
V. Results and Discussion

The following images will visually depict the process of our project.



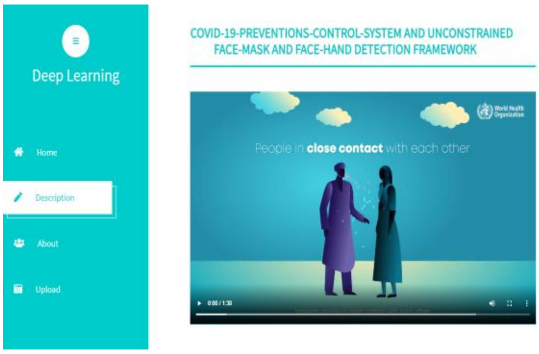
Home:

In this page will display the modules of a project



Description page:

In this page describe the over explanation about covid-19 virus.



About project page:

Here display the main theme of project.

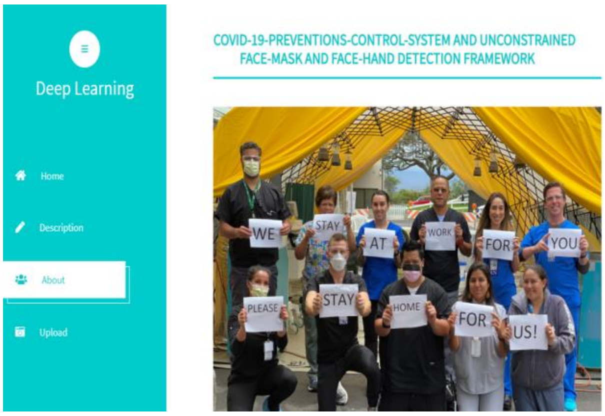
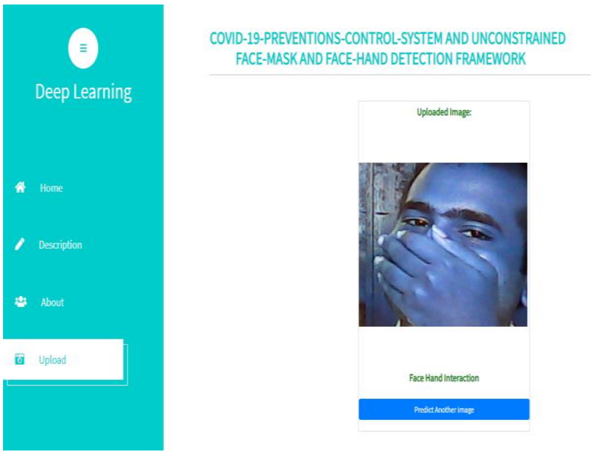


Image Upload page:

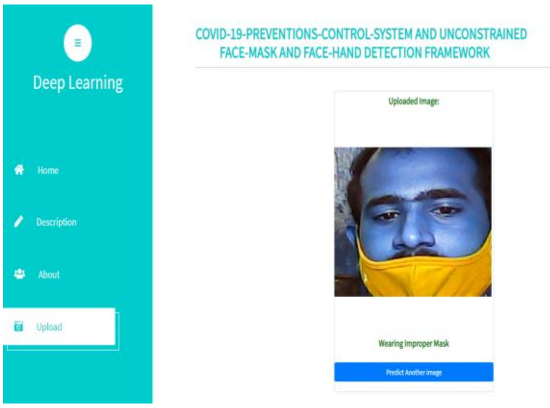
Here image files can be uploaded to classify whether the image which class it is.

View uploaded Image information:

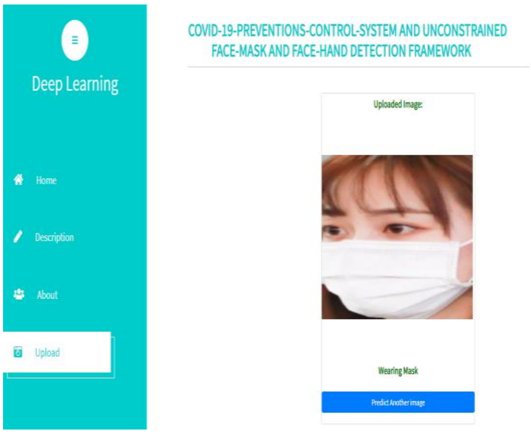
Here model classify the image is "Face Hand Interaction".



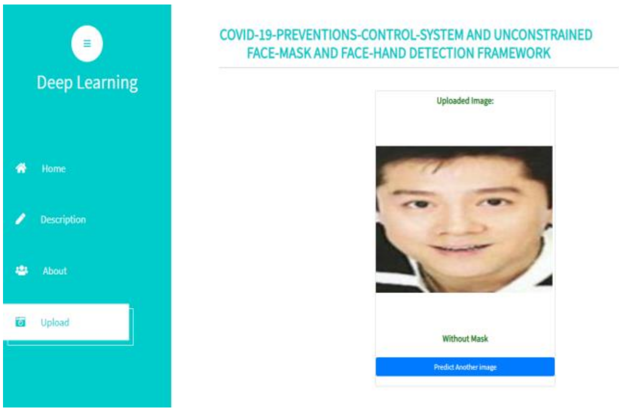
Here model classify the image is "Wearing Improper Mask".



Here model classify the image is "Wearing Mask".



Here model classify the image is “Without Mask”.



VI. Conclusion

In this project we have successfully developed a real time web application based image recognition system. This system detects people and it also detects if they are wearing a mask, improper mask, face hand interaction and without mask. In case they are wearing a mask, the type of mask wore by them is also recognized. To achieve this we have trained a pre-trained Mobile sNet model using facial images of people wearing masks. Computer vision technique are also used in order to process these images. We have observed a test accuracy of around 97%. In this system we form of bounding box over the recognized faces and the predicted classes with their probabilities are displayed over it.

VII. REFERENCES

[1]. Y. Fang, Y. Nie, and M. Penny, “Transmission dynamics of the covid-19 outbreak and effectiveness of government interventions: A datadriven analysis,”Journal of medical virology, vol. 92, no. 6, pp. 645–659, 2020.

[2]. S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, “Rational use of face masks in the covid19pandemic,”The Lancet Respiratory Medicine, 2020.

[3]. 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.

[4]. 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.

[5]. A. Kumar, A. Kaur, and M. Kumar, “Face detection techniques: a review,”Artificial Intelligence Review, 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, 2019.

Cite this article as :

Dr. S. Sathyalakshmi, Medikonda Venkata Leela Krishna, Phani Vineel Varma, Inkulu Sujith, Pelluri Venkata Sai Goutham Aditya, "COVID-19-Preventions-Control-System, Face-Mask, And Face-Hand Detection Framework", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 6, pp. 147-155, November-December 2022.
Journal URL : https://ijsrcseit.com/CSEIT228615