# ATTENDANCE SYSTEM USING FACE RECOGNITION
# AND
# MASK DETECTION

**R M NARESH KUMAR**

1934025aiml@cit.edu.in

**S SRIRAM**

1934048aiml@cit.edu.in

**S DEIVARANI**

deivarani@cit.edu.in

**Coimbatore Institute of Technology**

Peelamedu,

Coimbatore 641-014,

Tamilnadu, India

**Keywords**

Computer vision, COVID-19, Attendance, Mask, HOG, Haar-Classifier, Deep Learning, MobileNet, Face recognition, Face Detection, Face mask detection, Image processing.

---

## ABSTRACT

The goal of this system is to record attendance for the person (here it is for students) using face recognition and to detect the person who is wearing mask and who is not using Deep Learning. The recognition is done through camera. Dataset is images of persons to be recognized and persons with and without masks. Recognition is done through encoding calculation and comparison. These encodings are nothing but the feature matrix extracted from a specific portion [face] in the images. The algorithms used are haar-classifier and hog for face recognition and a deep learning model for mask detection. The mask-detection model is trained with a large set of images under 2 categories [With Mask, Without Mask] and it is a light weight model, which can also run in compact devices like mobile, tablets etc.

# 1. INTRODUCTION

For every organization, today attendance is the most important thing to record the presence of a person. The presence of a person in an organization is a sign that the person is carrying out their obligations to come to the agency or organization. Usually, attendance is done manually. It can be signed or called one by one. In this digital age, there must be a change from this absence to be able to accelerate and provide time efficiency. We can use face recognition to record attendance from everyone present in an organization. In this face recognition, many algorithms are performed to dissect and capture images of someone’s face, such as Machine Learning and Deep Learning. With this algorithm, the system can recognize a person’s face and record attendance from that person so that attendance activities are more efficient and faster.A face recognition system (FRS) is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source.

And about the mask-detection, nowadays it’s always recommended to wear a mask where ever we go in this pandemic situation, but there are some people who are being very lethargic and careless, who don’t wear a mask. In order to find them we have come up with detection model which detects the person who is wearing mask and who is not. Many precautionary measures have been taken to reduce the spread of this disease where wearing a mask is one of them. In this paper, we propose a system that restrict the growth of COVID-19 by finding out people who are not wearing any facial mask in a small organisation or school/college institution where all the public places are monitored with CCTV cameras.

It is hands free software that means just turn on the software, the other things will be taken care by the software itself. A deep learning architecture is trained on a dataset that consists of images of people with and without masks collected from kaggle. This program is in initial stage, we have also planned to update this software which enables the software to note the person’s name who are not wearing the mask and export that data to a xl sheet.

# 2. RELATIVE WORK

## ATTENDANCE SYSTEM

| S.No | ARTICLE | PUBLISHMENT | WORK |
|---|---|---|---|
| 1 | Attendance using Face Recognition | ISSN:2278-0181, Published by,www.ijert.org ENCADEMS - 2020 Conference Proceedings | Cascade classifier and LBPH (Local Binary Pattern Histogram) algorithms. |
| 2 | Attendance Marking System Using Face Recognition<br>By, Pooja L Kanth and Salva Biswal | Indian Journal of Science and Technology,<br>Year:2019, Volume: 12, Issue: 47, Pages: 1-3 | Done using PCA Analysis, Viola Jones Algorithm. |
| 3 | Real Time Automatic Attendance System for Face Recognition Using Face API and OpenCV<br>Sikandar Khan, Adeel Akram & Nighat Usman | Published in springer link on<br>31 March 2020,<br>Wireless communications:<br>**113**, pages469–480 | YOLO V3 (You only look once) algorithm for face detection and Microsoft Azure using face API for face recognition (face database). |

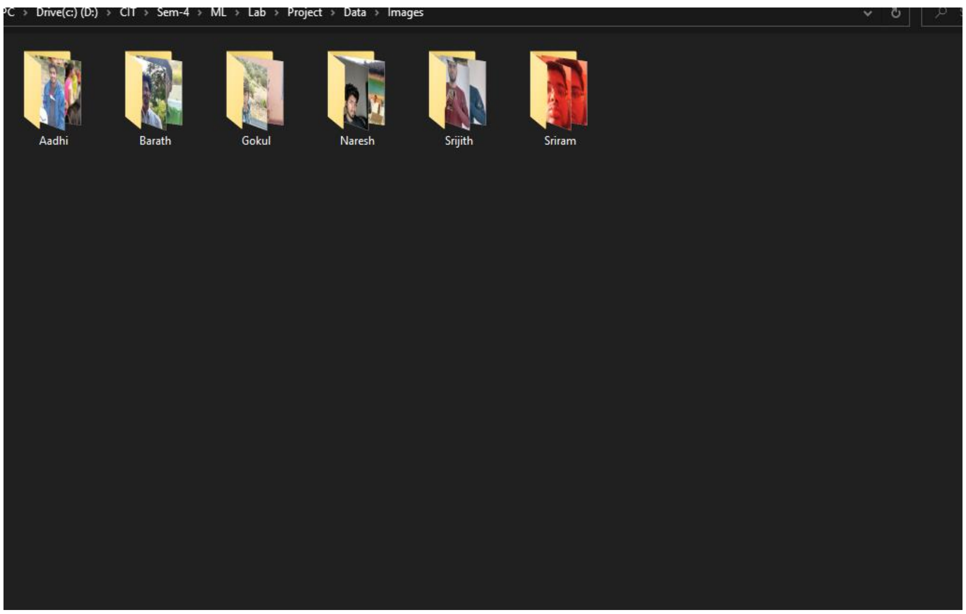
## MASK DETECTION

| S.No | ARTICLE | PUBLISHMENT | WORK |
|---|---|---|---|
| 1 | A real time DNN-based face mask detection system using single shot multi-box detector and MobileNetV2 | Sustain Cities Soc. 2021 Mar; 66: 102692.Published online 2020 Dec 31. | SSDMNV2 approach uses Single Shot Multi-box Detector as a face detector and MobilenetV2 architecture as a framework for the classifier. |
| 2 | AI on the Edge : Face Mask Detection - Sergio Virahonda | Published on Code Project<br>Date : 25 Jan 2021 | Combining Keras, MobileNet V2, and OpenCV to implement the model. |
| 3 | Face Mask Recognition using Machine Learning | IJIRT Paper ID: 150498<br>Published on : May 2021. | Using TensorFlow, Keras, Pytorch.<br>3 stages : Images Pre-processing – Detection – Classifier |

# 3. METHODS

### 3.1 Dataset Description

Dataset consists of images of various persons and around 1900+ images including both with/without mask [in a separate folder] and there are minimum 5 images of each person for attendance. These images are stored in the folders under the respective person names.




### 3.2 Data Pre-processing

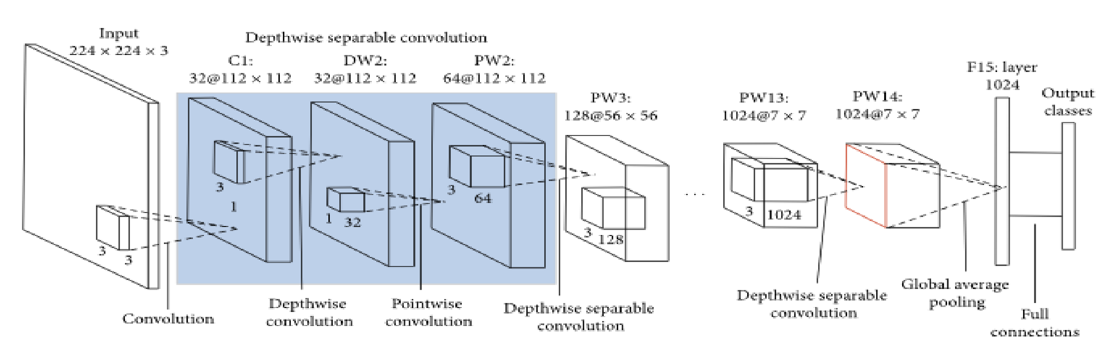
Data is based on images. So no major pre-processing needs to be done. Images are grey-scaled (black-white) and then face is detected for attendance but since we use a deep learning model (MobileNet) for mask detection, there isn't a need to convert the images into grey-scale, instead the images are processed as coloured image itself.

The image is resized into smaller size to reduce the load on the resources [GPU].

### 3.3 Exploratory Data Analysis


```
{'encodings': [array([-0.13925023,  0.1149301 ,  0.1259298 ,  0.01910496, -0.00343181,
       -0.06997918,  0.03061048, -0.04073902,  0.1516878 , -0.01173011,
        0.21591766, -0.04032761, -0.23574013, -0.13770893,  0.08768596,
        0.12724741, -0.10770105, -0.08220164,  0.01092249, -0.07982671,
        0.02984343,  0.00324398,  0.00843927,  0.16057235, -0.07547837,
       -0.38513121, -0.08522229, -0.14863114,  0.02721736, -0.09991058,
       -0.05025923,  0.07702903, -0.1090426 , -0.02370352, -0.0669633 ,
        0.07379948,  0.02556347, -0.09031876,  0.14632985,  0.05314435,
       -0.11069658, -0.09566651, -0.07692082,  0.26828685,  0.11348758,
        0.02619383,  0.08487093,  0.02584519,  0.0401193 , -0.17509958,
        0.00471109,  0.13902792,  0.07111607, -0.01813911,  0.02528663,
       -0.14589557, -0.03962337,  0.03140217, -0.25328058,  0.05260563,
       -0.00713092, -0.11280497, -0.10067859, -0.06983317,  0.21844296,
        0.1420919 , -0.10176162, -0.05064484,  0.19301541, -0.15393461,
       -0.03901505,  0.02221542, -0.12053358, -0.17044497, -0.22632432,
        0.12267815,  0.34299934,  0.14244539, -0.16664122,  0.02504197,
       -0.13000289, -0.01924819,  0.00887363,  0.09167656, -0.0962996 ,
        0.03898235, -0.10665497,  0.04128408,  0.07927088,  0.04486682,
        0.01266361,  0.2111609 , -0.03267941,  0.04351548,  0.00428436,
```


For attendance, the images are then used to find the encoding matrix. The face is detected in these images and then features are extracted using hog algorithm. The hog algorithm extracts features to get a normalized vector matrix. This is done for a small portion of the image. At last, the normalized vector covering all the portion of image is calculated which is the encoding matrix.




In mask detection, the MobileNet model uses depth wise separable convolutions. It significantly reduces the number of parameters when compared to the network with regular convolutions with the same depth in the nets. A convolutional layer is a matrix applied to images that perform a mathematical operation on the individual pixels to produce new pixels that are then passed as the input for the next layer and so on until the network's end is reached. The last layer is a single integer that turns the image output into a numerical class prediction that corresponds to an object we are trying to predict. This results in lightweight deep neural networks.

# 4. MODEL BUILDING

The Attendance system is built in 2 stages:

1. Detection
2. Recognition

## Detection

The detection is done using haar-classifier algorithm. The classifier for detecting face is built using xml file. The xml file is created using a GUI application (XMI file generator) or it can also be done using python source code by OpenCV functions. This classifier is also built-in in haar-cascade library. Using this, the face location tuple is calculated.

### Haar Classification algorithm

It is an Object Detection Algorithm used to identify faces in an image or a real time video. The algorithm uses edge or line detection features proposed by Viola and Jones in their research paper "Rapid Object Detection using a Boosted Cascade of Simple Features" published in 2001. The algorithm is given a lot of positive images consisting of faces, and a lot of negative images not consisting of any face to train on them.

## Recognition

The face-location tuple is used to extract features from the face located pixels. The HOG algorithm extracts the features from the pixels in the form of histograms. The histograms are formed using orientations. Gradients are used for the orientation calculation. The gradients are found using the changes in x and y directions. Then orientation is found using gradient values with the help of Pythagoras theorem [to determine angle]. These histograms are in the form of matrices. These matrices are then normalized to get the vector matrices which later form the encoding matrix for the face area in the image.

### Hog algorithm

HOG, or Histogram of Oriented Gradients, is a feature descriptor that is often used to extract features from image data. The HOG descriptor focuses on the structure or the shape of an object. Including the checking of edge is present or not (like normal descriptors), the HOG descriptor is able to provide the edge direction as well which is done by extracting the gradient and orientation (or you can say magnitude and direction) of the edges. Additionally, these orientations are calculated in 'localized' portions. This means that the complete image is broken down into smaller regions and for each region, the gradients and orientation are calculated. The HOG feature descriptor counts the occurrences of gradient orientation in localized portions of an image.

A loop is used to compare the live face encoding with the pre-stored encodings to find the probability of possibility of all the encodings. Then the encoding with higher probability is chosen, and the name included with the encoding is used to mark attendance.

Whereas, the mask-detection model is also built in 2 stages:

1. Detection (Face detection in real time).
2. Recognition (Recognising Mask on that face).

To reduce the stress on the resources, the following parameters are set.

**EPOCH -** It is set to 20. {Total Cycle}

**BATCH SIZE -** Is set to 32. {No of inputs per batch}

**LEARNING RATE -** Is set to 0.0001{for activation Function}

In our proposed model, the Depth Wise Separable Convolution is been used, in which the depth and spatial dimension of a filtered pixel matrix [from the detected portion of the image i.e., face] can be separated.

A depth wise separable convolution is made from two operations.

1. Depth wise convolution.
2. Point wise convolution.

## Depth wise convolution

- Depth wise convolution is the channel-wise DK×DK spatial convolution. Suppose in the figure above, and we have five channels; then, we will have 5 DK×DK spatial convolutions.
- It is a map of a single convolution on each input channel separately. Therefore its number of output channels is the same as the number of the input channels. Its computational cost is, **$Df^2 * M * Dk^2$.**

## Point wise convolution

- Point wise convolution is the 1×1 convolution to change the dimension.
- Convolution with a kernel size of 1x1 that simply combines the features created by the depth wise convolution. Its computational cost is, **$M * N * Df^2$.**

# 5. PERFORMANCE AND RESULTS

For the face recognition model, ROC and AUC curves are used to calculate the accuracy of each model for evaluation. The roc and auc curve is plot using the encoding matrix. The confusion matrix is found for this encoding matrix and from that the multi-class values is reduced to binary values using the average mid value (cut off).

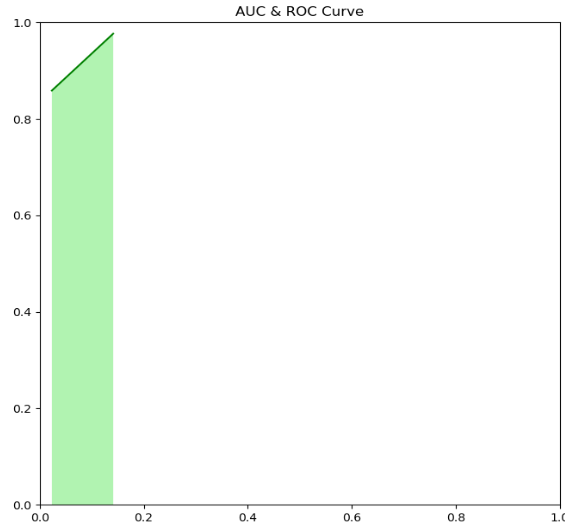
Haar- Classifier algorithm:

**Accuracy ~= 97%**

Hog algorithm:

**Accuracy = 0.8984375**

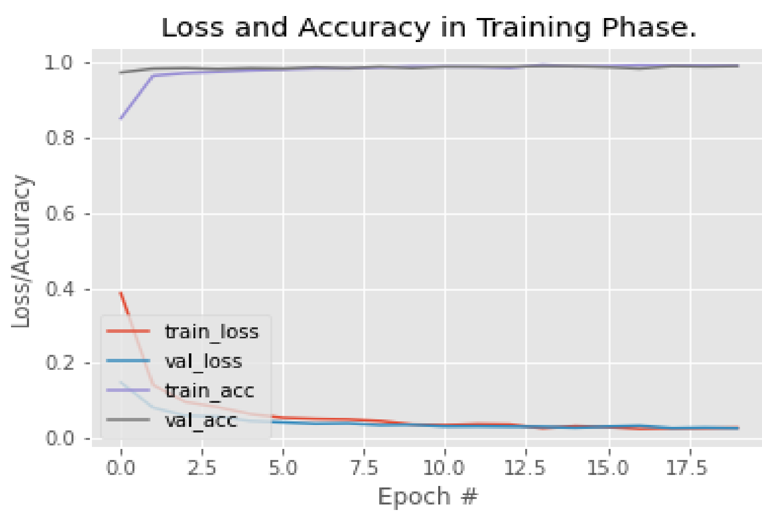
**Accuracy ~= 90%**



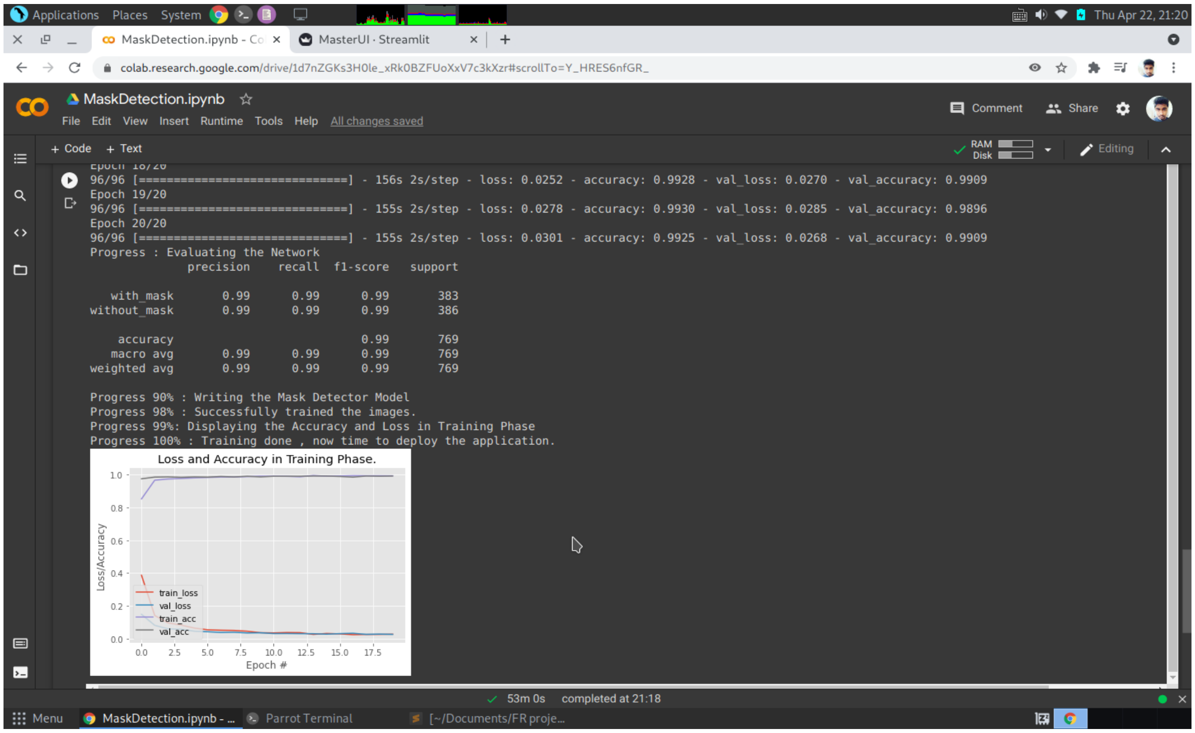

**HOG ALGORITHM**

In mask-detection model, the accuracy is measured through following means,

### Loss/accuracy graph



### Accuracy in training and testing phase




# 6. CONCLUSION

The proposed face-recognition model identifies the person's face and records the attendance and then for mask-detection, our proposed model is an efficient real-time deep learning model to detect the masked faces, where each face is identified in real-time with the help of bounding boxes. Attendance system is implemented using haar-classifier and hog algorithms whereas mask-detection is done through MobileNet. MobileNet Model is a light weight deep learning model and faster when compared to R-CNN, that's the reason we chose to work on that and this software can also be implemented in mobile phones and other portable devices too.

This Mask detection model gives around 85-90% accuracy and the Haar- Classifier and HOG algorithm has a accuracy of 97% and 90% for attendance system.

# 7. REFERENCES

## ATTENDANCE SYSTEM


- Hog Algorithm :
  - https://www.analyticsvidhya.com/blog/2019/09/feature-engineering-images-introduction-hog-feature-descriptor/

- Haar-Classifer:
  - https://towardsdatascience.com/face-detection-with-haar-cascade-727f68dafd08

Article References (S.no):

1. https://www.ijert.org/research/attendance-management-system-using-face-recognition-IJERTCONV8IS10012.pdf
2. https://indjst.org/articles/attendance-marking-system-using-face-recognition
3. https://link.springer.com/article/10.1007/s11277-020-07224-2


## MASK DETECTION


- MobileNet
  - https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470

- Article References (S.no)
  1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7775036/
  2. https://www.codeproject.com/Articles/5293068/AI-on-the-Edge-Face-Mask-Detection
  3. http://ijirt.org/Article?manuscript=150498

- Project Link
  - https://github.com/Nary-Vip/nary_ram_project.git