FACE MASK DETECTION

A SUMMER TRAINING REPORT

Submitted by

Arshiya Sarmai

19BCS6060

in partial fulfilment of Summer training for the award of the degree of

BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE
IN SPECIALISATION WITH
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

APEX INSTITUTE OF TECHNOLOGY
CHANDIGARH UNIVERSITY, GHARUAN, PUNJAB

JUNE-JULY 2021

ABOUT THE COMPANY

INTERNSHALA

Internshala is an internship and online based training platform and it is based in Gurugram, India. Internshala is a reputed platform to get numerous internship opportunities and various training courses. It was founded by Sarvesh Agrawal, an IIT Madras alumnus, in 2011. Internshala aims to provide students with relevant skills and practical exposure to get good start in their careers. It provide students with various opportunities by helping them to connect with people from different organisations in India. It gives students the exposure to evaluate themselves as to where they stand in terms of skill and how to improve them. They provide assistance regarding placements on how and what all things a student should focus on. It provide students with relevant skills and practical exposure to get good start in their careers.

CERTIFICATE

INTERNSHALA TRAININGS

Certificate of Training

Arshiya Sarmai

from Chandigarh University, has successfully completed a six weeks online training on **Deep Learning**. The training consisted of Getting Started with Deep Learning, Artificial Neural Networks, Convolutional Neural Networks, Image Recognition Project and Final Evaluation modules.

We wish Arshiya all the best for the future endeavours.

Sarvesh Agarwal
FOUNDER & CEO, INTERNSHALA

Date of certification: 2021-07-23 Certificate no. : 85C2FD84-B3C5-9DC0-7871-1A1B66D37CAF

For certificate authentication, please visit https://trainings.internshala.com/verify_certificate

ACKNOWLEDGEMENT

The training opportunity I had with Internshala was a great chance for learning and professional development. Therefore, I consider myself as a very lucky individual as I was provided with an opportunity to be a part of it.

Bearing in mind previous I am using this opportunity to express my deepest gratitude and special thanks to my teachers Abhishek Bansal and Pukhraj Parikh. I choose this moment to acknowledge their contribution gratefully. It is my radiant sentiment to place on record my best regards, deepest sense of gratitude for their careful and precious guidance which were extremely valuable for my study both theoretically and practically.

I perceive as this opportunity as a big milestone in my career development. I will strive to use gained skills and knowledge in the best possible way, and I will continue to work on their improvement, in order to attain desired career objectives.

I would like to thank Mr. Ashish Rastogi , Internship coordinator Department of CSE for their support. I am extremely grateful to my department staff members and friends who helped me in successful completion of this course.

Sincerely

Arshiya Sarmai

19BCS6060

ABSTRACT

Since we are amidst the pandemic and it is not very long since the first wave left us devastated and second wave is right here ,it becomes our prime objective to contribute in this scenario as per our capability.

There are certain precautions that one must take to avoid unwanted infections. Face mask and social distancing are two of the main precautions. Since most of the people are new with the concept of face masks, they are being irresponsible about wearing mask.

As our country starts going through various stages of reopening after the COVID-19 pandemic, World Health Organization (WHO) has declared the use of a face mask as a mandatory biosafety measure. Face masks are necessary in general public to curtail the spread of coronavirus. To mandate the use of facemask, it becomes essential to devise some technique that enforce individuals to apply a mask before exposure to public places.

This project focuses on identifying whether the person is wearing the mask (correctly) or not. In the present scenario due to Covid-19, there is no efficient face mask detection applications, but they are now in high demand for transportation areas, places for the social gatherings, large scale manufacturing areas and other enterprises to ensure safety.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In the recent times, the Corona virus have become very common, contagious and dangerous to the whole human kind. Since we are amidst the pandemic and it is not very long since the first wave left us devastated and second wave is right here ,it becomes our prime objective to contribute in this scenario as per our capability. The most affected and devastating condition is of India due to its extreme population in small area.

As our country starts going through various stages of reopening after the COVID-19 pandemic, World Health Organization has declared the use of a face mask as a mandatory biosafety measure. Face masks are necessary in general public to curtail the spread of coronavirus. To mandate the use of facemask, it becomes essential to devise some technique that enforce individuals to apply a mask before exposure to public places.

This project focuses on identifying whether the person is wearing the mask (correctly) or not. In the present scenario due to Covid-19, there is no efficient face mask detection applications, but they are now in high demand for transportation areas, places for the social gatherings, large scale manufacturing areas and other enterprises to ensure safety.
We have used convolutional neural network for the same. The model is trained on a real world dataset and tested with live video streaming with a good accuracy. Further the accuracy of the model with different hyper parameters and multiple people at different distance and location of the frame is done.

THEORY

EXISTING SYSTEM

In [1] they have proposed a pre-trained MobileNet with a global pooling block for face mask detection. The pre- prepared MobileNet takes a shading picture and creates a multi-dimensional component map. The worldwide pooling block that has been used in the proposed model changes the element map into an element vector of 64 highlights. At long last, the softmax layer performs paired order utilizing the 64 highlights.

In [2] paper utilizes a proficient and strong item location calculation to naturally identify the appearances with veils or without covers, making the plague avoidance work more clever. In particular, they gathered a broad data set of 9886 pictures of individuals with and without face covers and physically named them, at that point use multi-scale preparing and picture mistake techniques to improve YOLOv3, an article recognition calculation, to consequently distinguish whether a face is wearing a veil.

Shervin Emami[3] in his paper proposed a solution to the problem of face recognition using OpenCV. His proposed system is Eigenfaces using grayscale images. His paper demonstrated a process to convert color images to grayscale and then applied histogram Equalization to it as a very simple method of automatically standardizing the brightness and contrast of facial images. His paper concluded the various ways in which the accuracy of image recognition system can be improved.

Faizan Ahmad et al.[4] in their paper, presented a “State of the Art” face detection and recognition system. Their main work is focused on video based face detection and recognition surveillance. Initially they have used Adaboost and SVM algorithm to find the accuracy. In comparison to SVM, Haar Cascade classifier of Adaboost showed greater mean accuracy(96.70%). Thereafter, they have used PCA, LDA, LBP and Gabor feature based model for face detection and recognition on five dataset and presented a comparison of accuracy. According to the results presented by them, Gabor filer achieved a mean accuracy of 92.35% which is greater then the mean accuracy of rest of the algorithms. PCA have achieved the lowest mean accuracy among the others. They have concluded that Gabor is better as it’s qualities overcomes the data sets’ complexity.

M. Naveenkumar and A.Vadivel[5] have discussed about sample real time image processing applications of OpenCV are along with steps. They have presented an edge detection system using OpenCV for image detection.

Ashu Kumar et al.[6], in their article ,have presented survey of various face-detection techniques. They have discussed about the different techniques for instance feature based model and image based approaches. They have also presented a comparative study of the two methods. A slightly detail study of various methods in each of the above approaches such as HMM, SVM, Neural Networks and Adaboost have also been presented in the paper. They have concluded the paper by providing information about various standard datasets for face detection.

PROPOSED SYSTEM

The model proposed here is designed and modeled using python libraries namely Tensorflow, Keras and OpenCV. The model we used is the VGG19of convolutional neural network. The method of using VGG19 is called using Transfer Learning. Transfer learning is using some pre trained model to train your present model and get the prediction which saves time and makes using training the different models easy. We tune the model with the hyper parameters learning rate, number of epochs and batch size.

A dataset is used to train the models. The dataset consists of 12,000 images belonging to 2 classes: Masks and No masks. Model preprocessing is done on the images and then CNN models are used to predict the model.

METHODOLOGY

The methodology used in this project has been depicted in the figure below.

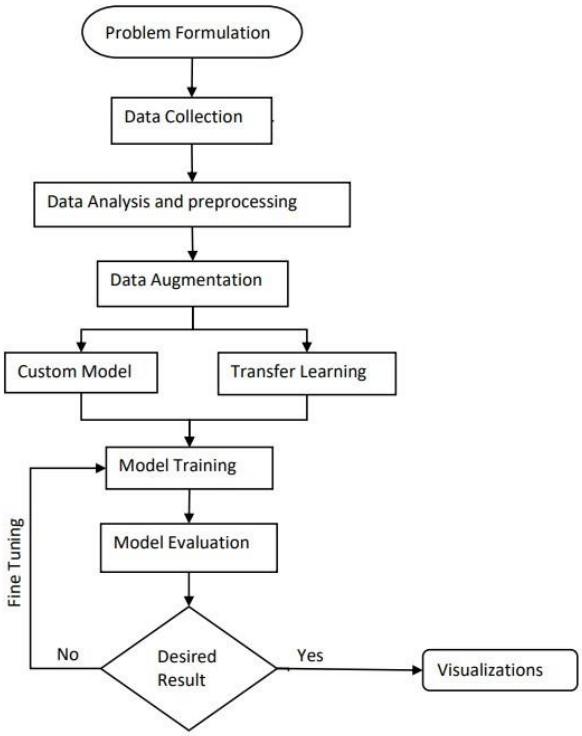

Figure 1

The methodology includes the preprocessing of the dataset in order to train the CNN models. Data preprocessing involves conversion of data from a given format to much more user friendly, desired and meaningful format. It can be in any form like tables, images, videos, graphs, etc. These organized information fit in with an information model or composition and captures relationship between different entities [6]. The proposed method deals with image and video data using Numpy and OpenCV.

Detailed methodology used in this project is presented below.

The Dataset:

The dataset contained 12,000 images of 2 classes, masks and no masks. The images in the dataset are in RGB format. There are 10,000 images in training set, 1,000 images in each of validation and test set.

The dataset is a balanced dataset. So accuracy metrics is sufficient to predict the performance of the CNN models.

A sample of the images in the dataset is presented in figure below.

Figure 2

Figure 3

The images in the mentioned datasets are in RGB format with pixels values in the range [0,255]. This pose a problem while initializing weights in the neurons of the CNN model. Without scaling, the high pixel range images will have large amount of votes to determine how to update weights. We rescaled the images from [0,255] to [0,1] using a rescaling factor

of 1/255. Scaling every images to the same range [0,1] will make images contributes more evenly to the total loss.

The number of images in the dataset is not sufficient to train an image classifier to predict masks accurately in real time. Image augmentation is used to artificially increase the size of the datasets in order to generalize the prediction and increase the prediction accuracy of new unseen data. The existing images are slightly altered in order to create more images to train the DCNN models on. The augmentation. The augmentation parameters used during training of the images are mentioned in table 1.

Table 1
Using data augmentation parameters and values.

Data Augmentation Parameter	Parameter Value
Rotation_range	20
Height_shift_range	0.2
Width_shift_range	0.2
Shear_range	0.2
Zoom_range	0.2
Vertical_flip	True
Horizontal_flip	True

Table 1

Figure 4

The size of images for training the DCNN models is changed based on the size of images in the dataset. They are changed to 128×128×3.

The model is implemented using Keras sequential API. It is constructed to predict one of the following two predictions in the dataset: Mask and No mask. The model contains Convolution layers, MaxPooling layers, Dropout layers and fully connected layers.

The CNN model consists of only 557,183 parameters. The input shape of the model indicates that the breast histopathology images used for training are of shape 128×128×3 having all three additive colour spaces i.e. Red, green and blue. The output has a shape of 2 for the classification of 2 sample classes in the dataset. The small numbers of parameters of the model indicate that the proposed model is lightweight.

The proposed architecture comprises a heterogeneous mix of convolution layers with a kernel size of 3×3 and pooling layers of pool size 2×2. A varied filter size for the dimensionality of the output space is used in different convolution layers of the architecture. This architectural diversity of the CNN architecture provides a strengthened feature detection from the input mask images and enhances the representational capacity of the task of mask detection.

ReLU activation function is used in the hidden layers and sigmoid activation function is used in the output layer of the model.

VGG19: VGG19 is a variant of VGG model which in short consists of 19 layers (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). There are other variants of VGG like VGG11, VGG16 and others. VGG19 has **19.6 billion FLOPs**.

Properties of VGG19:

- A fixed size of (224 * 224) RGB image was given as input to this network which means that the matrix was of shape (224,224,3).
- The only pre-processing that was done is that they subtracted the mean RGB value from each pixel, computed over the whole training set.
- Used kernels of (3 * 3) size with a stride size of 1 pixel, this enabled them to cover the whole notion of the image.
- Spatial padding was used to preserve the spatial resolution of the image.
- Max pooling was performed over a 2 * 2 pixel windows with stride 2.
- This was followed by Rectified linear unit(ReLU) to introduce non-linearity to make the model classify better and to improve computational time as the previous models used tanh or sigmoid functions this proved much better than those.
- Implemented three fully connected layers from which first two were of size 4096 and after that a layer with 1000 channels for 1000-way *ILSVRC* classification and the final layer is a softmax function.

RESULTS AND DISCUSSIONS

The model achieved an accuracy of 99.75% on the training set and 99.19% on the test set. This shows that the model is predicting the classes excellently. The training loss is 0.0065 and test loss is 0.013. The training and validation accuracy and loss of the model is presented in figure 5.

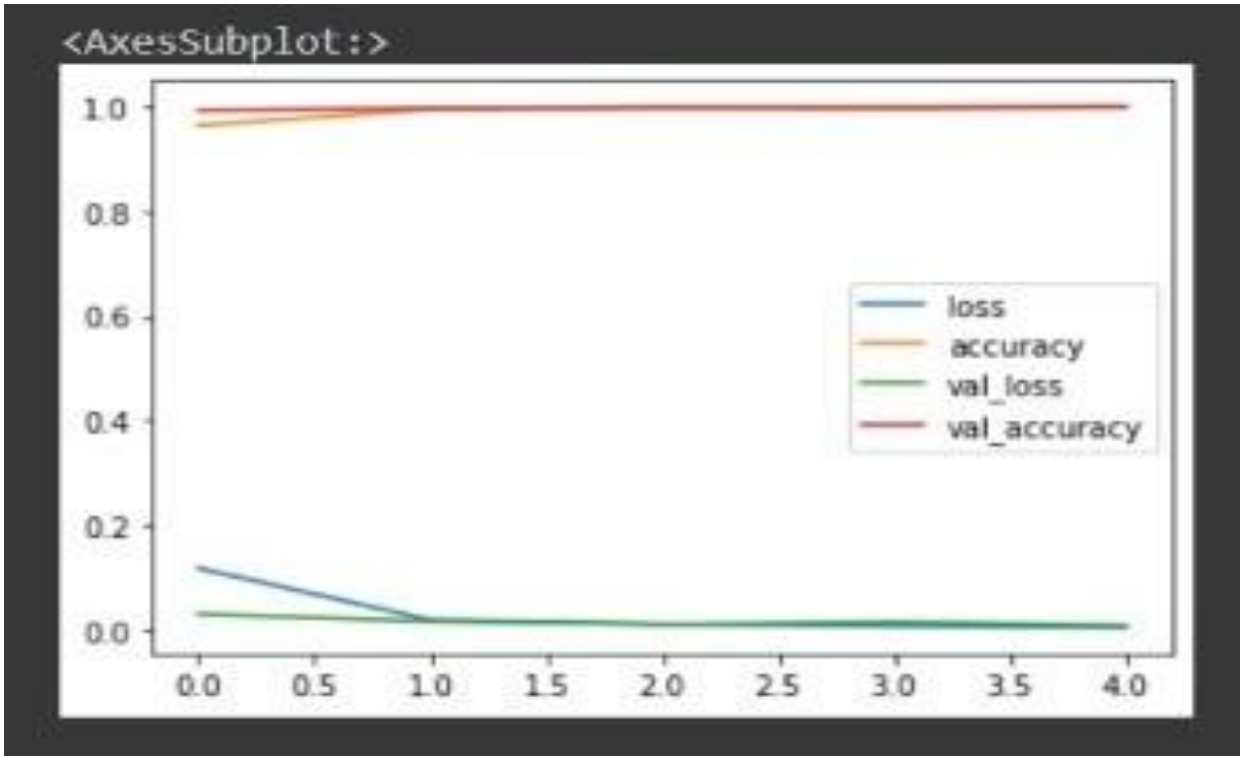

Figure 5

It is clear from the figure that the accuracy of the model almost converges at the 10^{th} epoch which shows the model is performing very good.

The comparison of the two model is presented in table 2.

	Classification Accuracy		**Loss Value**	
Models	**Training set**	**Test set**	**Training set**	**Test set**
VGG19	99.75%	99.19%	0.0065	0.0135

Table 2

VGG19 performs good on the test set. There is no overfitting in the model. It signifies that the model is able to generalize the dataset.

CONCLUSIONS AND FUTURE SCOPE

In this project , a deep learning-based approach for detecting masks over faces in public places to curtail the community spread of Coronavirus is presented. This application can be used at various places such as transportation areas, places for the social gatherings, large scale manufacturing areas and other enterprises to ensure safety.

The mask detection model has achieved an excellent training accuracy of 99.2%.

The present model proposed gives great accuracy for single face with and without mask. For multiple faces also it gives quite good accuracy. It works easily on any mobile device just by switching on the video stream, with no external hardware requirement. Further we will work for improving the accuracy for multiple face mask detection, to classify the faces into three categories that is, With mask, without mask, Improper mask instead of just the two with and without mask class by adding datasets with images of people wearing masks not covering their noses properly. This project will have a huge application given the situation of Covid-19 pandemic. This mask- no mask surveillance can be done remotely without human interference which can be very helpful.

REFERENCES:

1. [1] A. G. Howard, M. Zhu, B. Chen et al., “Mobilenets: efficient convolutional neural networks for mobile vision applications,” 2017, https://arxiv.org/abs/1704.04861.
2. [2] Wei Wang, Yutao Li, Ting Zou, Xin Wang, Jieyu You, Yanhong Luo, "A Novel Image Classification Approach via Dense-MobileNet Models", Mobile Information Systems, vol. 2020,ArticleID 7602384, 8 pages, 2020. https://doi.org/10.11 55/2020/7602384
3. [3] I. B. Venkateswarlu, J. Kakarla and S. Prakash, "Face mask detection using MobileNet and Global Pooling Block," 4 2020 IEEE 4th Conference on Information & Communication Technology (CICT), 2020, pp. 1-5, doi: 10.1109/CICT51604.2020.9312083.
4. [4] M. S. Ejaz and M. R. Islam, "Masked Face Recognition Using Convolutional Neural Network," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019, pp. 1-6, doi: 10.1109/STI47673.2019.9068044
5. [5] Changjin Li, Jian Cao, and Xing Zhang. 2020. Robust Deep Learning Method to Detect Face Masks. In *Proceedings of the 2nd International Conference on Artificial Intelligence and Advanced Manufacture* (*AIAM2020*). Association for Computing Machinery, New York, NY, USA, 74–77. DOI:https://doi.org/10.1145/3421766.3421768
6. https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset