FACE MASK DETECTION USING CNN

Parth Sourav
Student of CSE department
MANIT Bhopal
Souravparth26@gmail.com

Anubhav Goyal
Student of CSE department
MANIT Bhopal
anubhavgoyal30688@gmail.com

Under the guidance of

MS. Priyanka Verma mam
MANIT Bhopal
303priyanka.verma@gmail.com

Abstract-Face mask detection has become a obvious need in this covid era.This research paper proposed an approach which is very efficient and fast in detecting face mask.The work has been completed in the following steps respectively,image preprocessing,training,classification.

Keywords-CNN, image classification

1.INTRODUCTION

[1]Coronavirus disease 2019 (COVID-19) is defined as illness caused by a novel coronavirus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; formerly called 2019-nCoV), which was first identified amid an outbreak of respiratory illness cases in Wuhan City, Hubei Province, China.[1] It was initially reported to the World Health Organization (WHO) on December 31, 2019. On January 30, 2020, the WHO declared the COVID-19 outbreak a global health emergency.

[2]4,145,699 people have died so far from the coronavirus COVID-19 outbreak as of July 22, 2021.

Although most of the countries has started doing vaccination on a large scale ,but it is not guaranteed the full prevention from covid-19. In this case prevention and awareness has become equally important along with vaccination.Mask wearing and sanitizing hands regularly has become very important .

In public places like malls,offices,educational institutions it is very difficult to monitor each individual ,so we are to develop an artificial mask recognition system ,by which we can detect whether a person is wearing mask or not.

In our model we are using CNN architecture as a feature extractor and dense neural network as classifier.

2. Related works

In this paper our main focus is on finding out finding out whether the person is wearing a mask or not.

In [3] authors achieved 93% accuracy by using YOLOV3 algorithm for face detection.It was trained on dataset CELEBA and WIDERFACE.

Sheikh [4] used deep learning model with VGG-16 classifier .It was trained on KDEF dataset and it achieved 88% accuracy.

In [5] a boosting based classification,boosted cascades with haar features were embraced using Viola -Jones detector, then a Multiview face mask detector was made.A face Mask detector model was also made using decision tree algorithm .

Zhang, Zhang,Li, and Qiao (2016) [6],made a DPM(Deformable Part Model) based model.In DPM based classification, the structure and orientation of several different faces are modelled. Ramanam proposed a random forest tree model in face mask detection which guesses face structures

and face poses. DPM based model can be very accurate but its cost of computation is very high.

Z. Abidin [7] discussed facial expression recognitionusing fisherface, where the accuracy achieved 89.20%.

P.Pattanasethanon et al. [8] proposed a system which recognizea face with the help of eigenface. They also focused on special regions of the face as nostril areas and oral areas, although they had a little amount of dataset, they achieved 100%accuracy for their work.

T. Schenke et al. [9] described a facial recognition system with the help of raspberry pi using CNN, KNN with eigenface, and SVM, where the SVM algorithm showed higher accuracy for facial recognition.

There exist some limitation in these works some are not so accurate ,some have very high computational power and some are trained on a limited dataset. Hence we are to propsed an efficient approach.

3.Proposed model

This project used CNN model to detect mask ,which is giving quite significant accuracy. It takes an RGB input image from any orientation to obtaining output. The main work of this function is feature extraction and class prediction to the images. In the feature extraction system, the image is sketched and created into a new image where the generated image is more efficient than the previous image.

We are using convolutional layer for feature extraction which consists of the operation like max pooling and,activation functions.Max pooling operation is used to extract the high intensity pixels from the image which is the important feature.Convolutional layer after feature extraction gives a set of images with labels as output for better prediction.

The input collection of images should be preprocessed before feature extraction.Preprocessing consists of operation like reducing the size ,removal of noises etc. so that CNN layer can extract the feature efficiently.

The project involves a series of steps like image preprocessing,feature extraction ,image classification and finally image prediction.

To detect a mask on face first of all we need to extract face from image that we captures,for this we are using a haar cascade classifier for frontal face extraction which is [10] haarcascade_frontalface_alt_tree.xml from github.

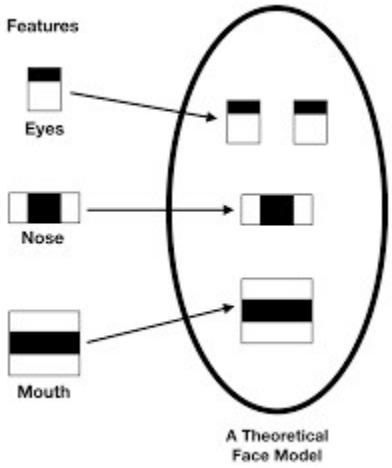


Fig:Cascade classifier face extraction

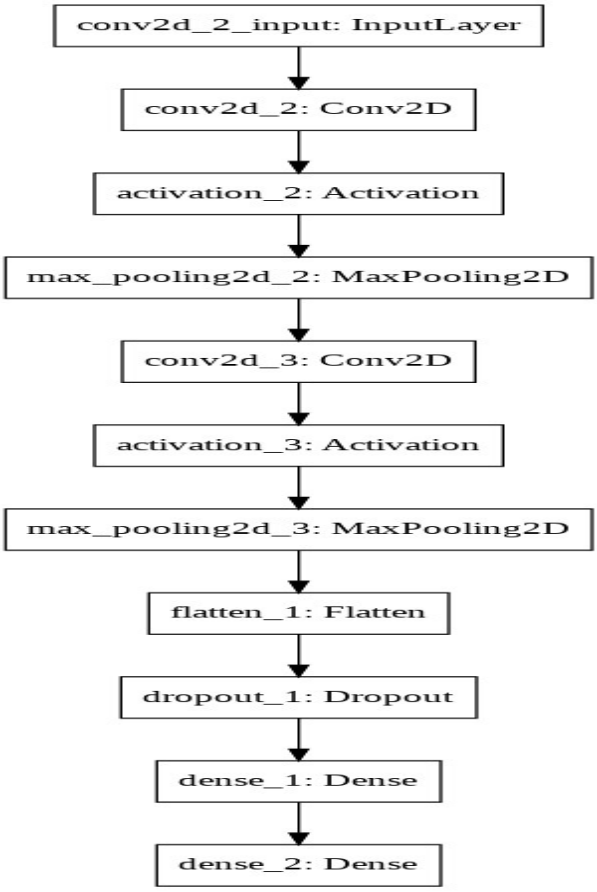


Fig: CNN architecture

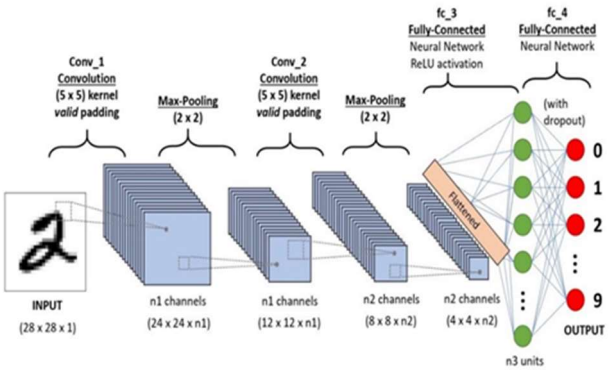
3.a Data set used

The artificial dataset created by Prajna Bhandary took standard im-ages of faces and applied facial landmarks. Facial landmarks allowed to locate facial features of a person like eyes, eyebrows, nose, mouth, and jawline. This used an artificial way to create a dataset by including a mask on a non-masked person image. Still, those images were not again used in the artificial generation process. The use of non-face mask samples involved the risk of the model becoming heavily biased. It was a risk to use such dataset images from various other sources. So they have included a dataset which would consist of images of people with masks and unmasked, which compensated for the error correction.

In the end, a dataset which included 5521 images was obtained,having the label "with_mask" and "without_mask" also contained 5521 images to make a balanced dataset. The distribution between the two classes is visualized in Fig. 2. The created dataset can also be used for detecting assailants who cover their faces while performing unlawful deeds. The dataset has been made available at

https://github.com/TheSSJ2612/Real-Time-Medical-Mask-Detection/releases/download/v0.1/Dataset.zip.

3.b. Experimental Analysis



The model consist of 3 parts -:

1)Image Recognition -:

This part can be summed to convolution +pooling+activation layer.

Convolution

A filter is a matrix which extracts features from an image (this is where the learning takes place). The dot product between the filter values and the image pixel values is forms the convolution layer.

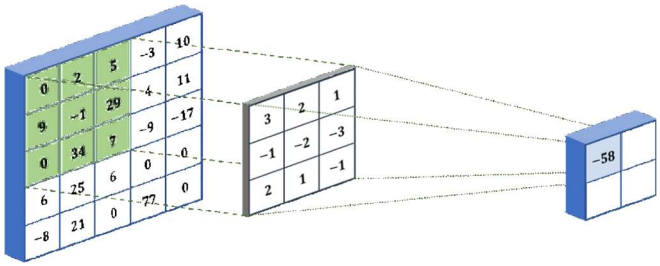


Fig:covolution

$$C(T) = (A * B)(x) = \int A(T) \times B(T - x)\, dT$$

Covoltion operation between two matrices

Pooling

Pooling is done to reduce the size of image so that our model can be trained faster. It also deal with the problem of overfitting by dropping unwanted variables.

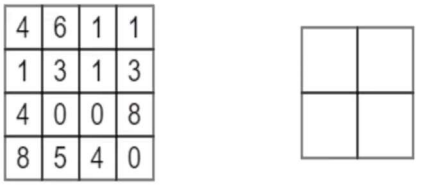


Fig: pooling

Activation

In our case we are takin ReLu activation function which gives positive outcomes same as the input for positive values but for negative values gives '0 as output.

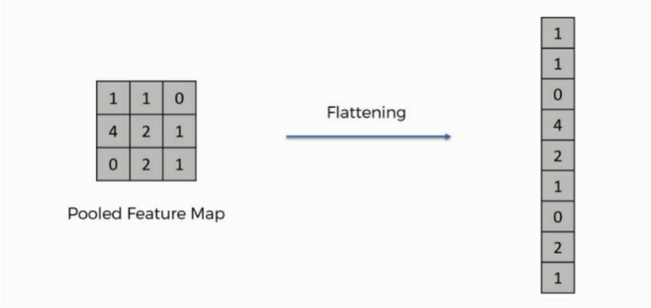
Relu : f(x) = max(0, x)

2)Image classification-:

This is a 'fully connected neuron' network working as a classifier

Flattening

Flattening is necessary before entering the convoluted input in the dense layer.Classification is easy if the tensor is in 1d array.



Dropout

Dropping out neurons means deactivating some neurons to speed up the the process of training and back propogation.

3.c Steps and pseudocode

The basic steps to build an image classification model using a neural network are:

=> Normalize the image pixel values (divide by 255):

```
images=np.array(images)/255.0
```

=> One-Hot Encode the categorical column:

```
lb=LabelBinarizer(
labels=lb.fit_transform(labels)
labels=to_categorical(labels)
labels=np.array(labels)
```

=> Build a model architecture (Sequential) with Dense layers:

```
model.add(Dense(64,activation='relu'))
model.add(Dense(num_classes,activation='softmax'))
```

=>Train the model and make predictions

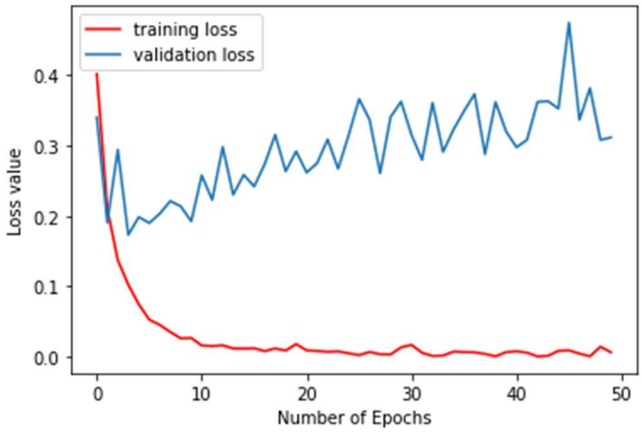
```
model.compile(loss='categorical_crossentropy',
optimizer=Adam(lr=0.001),
metrics=['accuracy'])
fitted_model=model.fit(
train_X,
train_y,
epochs=epochs,
validation_split=0.25)
```

5. Results

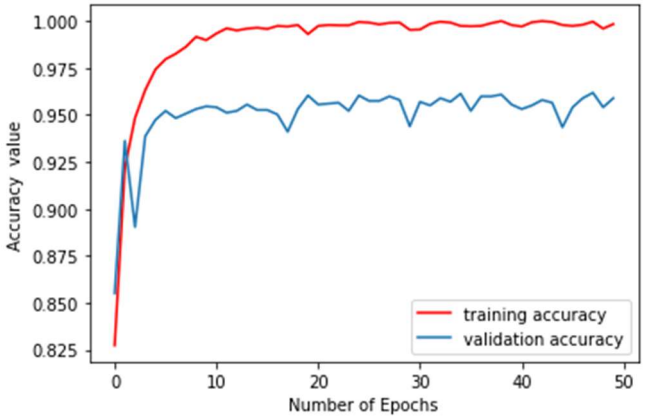
The model is giving Quite high accuracy of 99% upto last epoch. We have divided the dataset into two parts 75% of dataset for training and 25% for validation.

On an average the model is giving 97% of accuracy.

loss graph-:



Accuracy curve-:



The classification matrix formulae are given by

Accuracy = Tp + Tn/(Tp + Fp + Fn + Tn)

Precision = Tp/(Tp + Fn)

Recall = Tp/(Tp + Fn)

f 1 score = 2 ∗ Recall ∗ Precision/(Recall + Precision)

Where Tp = True positive,

Tn = True negative,

Fp = False positive,

Fn = False negative

In our case all parameters we obtain are nearly 96%.

6. Future Works.

Model is less sensitive to the skin colour masks,transparent masks and the animal wearing masks ,so in future model can be trained by taking these things into consideration.

An alert system can be developed so that we can get notification whenever alert system encounters a person who is not wearing masks.

7.Conclusion

This paper is based on the establishment of a mask monitoring system using efficient CNN model.The data set by Prajna Bhandary has been selected very smartly and hence helpful in training. It worked accurately and gave notable results.

Reference

[1] World Health Organization. Prevention, identification and management of health worker infection in the context of COVID-19. 2020

[2] "WHO Coronavirus Disease (COVID-19)

Dashboard." https://covid19.who.int/ (accessed May 21, 2020).

[3]C. Li, R. Wang, J. Li, L. Fei, Face detection based on YOLOv3, in:: Recent Trends inIntelligent Computing, Communication and Devices, Singapore, 2020, pp.277–284, doi: 10.1007/978-981-13-9406-5_34.

[4] S. A. Hussain, A.S.A.A. Balushi, A real time face emotion classification andrecognition using deep learning model, J. Phys.: Conf. Ser. 1432 (2020) 012087,doi: 10.1088/1742-6596/1432/1/012087.

[5] Jones P. Viola P. and Jones M.(2001).Rapid object detection using a boosted cascade of simple feature,Charles Rich:University of Rochester.

[6] Zhang, Zhang,Li, and Qiao (2016),

[7] Z. Abidin, A. Harjoko, "A Neural Network based Facial ExpressionRecognition using Fisherface", International Journal of Computer Applications (0975 – 8887), Vol.59–No.3, 2012

[8] P. Pattana sethanon, C. Savithi," Human Face Detection and Recognition using Web-Cam", Journal of Computer Science 8 (9): 1585-1593, 2012

[9] T. Schenkel, O. Ringhage, N. Branding," A COMPARATIVE STUDYOF FACIAL RECOGNITION TECHNIQUES With focus on low computational power", 2019

[10] https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_frontalface_default.xml