Rakshak - PRML Minor Project
Detecting whether the person is wearing mask or not!

Vikram (B19CSE098) and Nikhil Raj(B19CSE059)

Abstract

In this documentation, we presented all the required information for deciding the best machine learning model to determine if a person is wearing mask or not. Apart from this we have done all the necessary visualization using different plots using various inbuilt functions from the sklearn library.

Index Terms

1) Introduction
2) Data Visualization
3) Models Used
4) Observation
5) Conclusion

I. Introduction

The mask or no mask project comes under supervised classification problem. In this problem statement we were expected to classify if the person in the given image is wearing masks or not. However, masks can be of any colour, any size and any type. Perhaps, considering all the factors, high accuracy and precision was expected from us.

For this problem statement we have been asked to use three different classifiers as taught in the PRML course with best hyper parameters to improve the overall accuracy. The three models that were used are Convolutional Neural Network, Random Forest Classifier, Support Vector Machine Classifer whose explanation will be formally provided in the MODELs USED section in the document.

II. Data Visualisation and preprocessing

In this section all the plots that, that we found relevant subjected to the dataset are plotted. Data provided in dataset are of different shapes. All provided images are rescaled to (128,128,3).
After scaling of images, data is converted to numpy array, for further use in trainig and testing models. For normalisation, All values in dataset are divided by 255, so that to confine the data in range of [0,1]. Now, Dataset in form of a numpy array of 3d-matrices. Other than that there is no such recognizable trend or pattern in dataset.

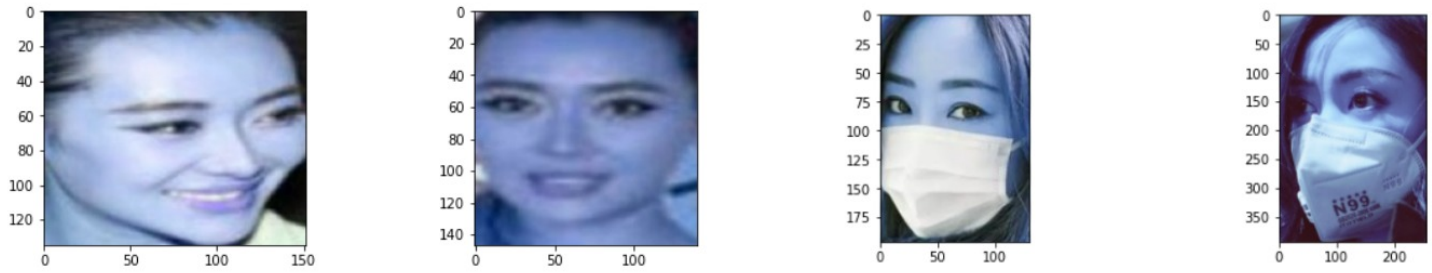

Fig. 1. After rescaling, sample of some data are plotted with First two images with First two without mask and last two with mask.

Data used for Model training had very imbalanced data with around 40,000 images with out mask and only 2000 images with mask. When Convolutional Neural Network model was trained over 50 percent of the data it showed a accuracy of 98 percent after 10 epochs. But the model was very miss leading. It is predicting all images as without mask.

For tackling over this situation No-mask images were reduced to 2500 to make a valid dataset. It reduces the size of training set by a huge amount but for good.

After reshaping the data, dataset contains a total of 4638 images with 2500 without mask and rest 2138 with mask. Which is a almost fare distribution.

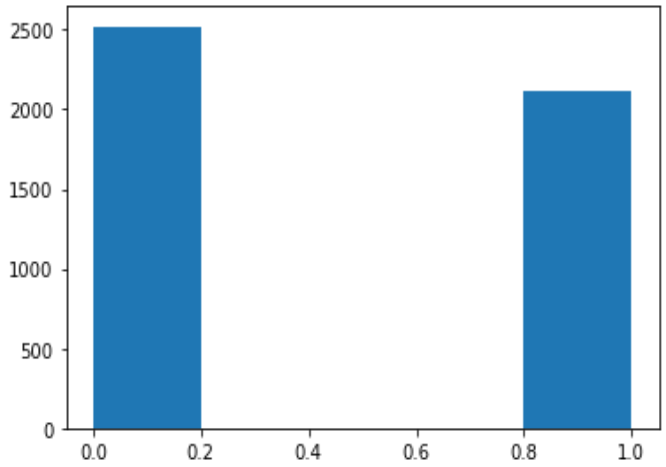

Fig. 2. Plot of count of each class present in working dataset. 0 implies a no-mask image and 1 implies image having atleast a mask.

III. Models Used

1) Random Forest Classifier :
   Random Forest can be said as a subset of Boosting algorithm. In the context of mask vs No mask problem, Random Forest uses voting technique to determine the final class from the attribute given. So Random Forest Algorithm choose the parameters which ensure the presence of Glass or No Glass, and classifies according to that.
2) Support Vector Machine:
   Support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. In our dataset, each marked for belonging to one of two categories 0 and 1, the SVM classifier builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. Our SVM model represents the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.
3) Convolutional neural network:
   In CNN, what is happening is that, it takes all the values of each pixel, which consists of features of the image and fires neurons corresponding to every new hidden layers formed. Based on the initial weights it classifies the images giving labels 0 and 1 representing persons wearing no-mask and mask respectively and if the classification is wrong back propagation takes place where the weights are recomputed and the same process takes place.
   In this project, Sequential CNN model by keras is used, activation function used is relu function, with 4 layer in network.

IV. Observations

1) Model:

```
Model: "model_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_4_input (InputLayer)  [(None, 128, 128, 3)]     0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 128, 128, 32)      896
_________________________________________________________________
batch_normalization_4 (Batch (None, 128, 128, 32)      128
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 128, 128, 32)      9248
_________________________________________________________________
batch_normalization_5 (Batch (None, 128, 128, 32)      128
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 64, 64, 32)        0
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 64, 64, 64)        18496
_________________________________________________________________
batch_normalization_6 (Batch (None, 64, 64, 64)        256
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 64, 64, 64)        36928
_________________________________________________________________
batch_normalization_7 (Batch (None, 64, 64, 64)        256
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 32, 32, 64)        0
_________________________________________________________________
flatten_1 (Flatten)          (None, 65536)             0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               8388736
_________________________________________________________________
dense_3 (Dense)              (None, 2)                 258
=================================================================
Total params: 8,455,330
Trainable params: 8,454,946
Non-trainable params: 384
_________________________________________________________________
```

Fig. 3. CNN with output format of each layer and number of parameters for each layer.

Only 5 epochs are used in CNN, more than that overfitting the data, for 6-8 epochs loss is not changing and after that loss function starts to rise high.
Loss value on validation dataset after each epoch is: 3.7331, 2.7875, 2.0753, 0.2372, 0.4579 Accuracy on validation set after each epoch: 0.9228, 0.9340, 0.9539, 0.9832, 0.9815

Fig. 4. Loss trend train and valid dataset before reducing number of data.

2) Efficiency of all three classfiers :
Random Forest Classifier gives a score of 99.05 percent.

SVM Classifier gives a score of 98.16 percent.

CNN gives a score of 98.2 percent.

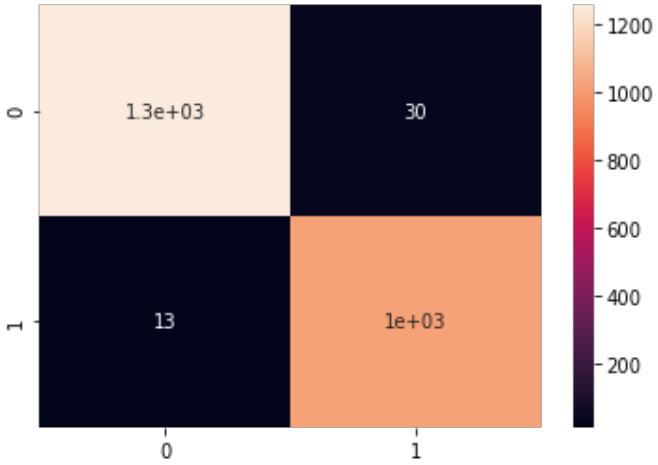

Fig. 5. Heatmap showing true positive, true negative, false negative, true negative.

Recall for these predictions is 0.990 and F1-score is 0.9834.
Only 43 false predictions out of 2300 cases is quite a good accuracy for neural network on training dataset of only 2300. But here, SVM classifier and Random forest both performed better than CNN.
Reason for that is small dataset for training only 2300 data-points which quite insufficient to train a neural network. SVM uses to find hyper plane in an n-dimensional space due to very high dimensions of original data, SVM may have been outrunned by Random forest. So,Random Forest outperformed CNN and SVM classifier both.

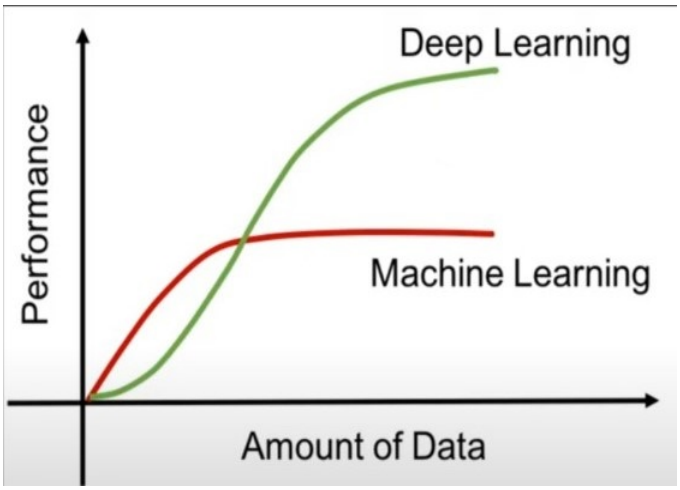

Fig. 6. Graph here shows that for smaller datasets Machine learning models performs better over Deep Learning models.

V. Conclusion

Among all three classifiers used, Random Forest Classifier performed better than CNN and SVM and gives an exceptional accuracy score of 99.05 percent and hence using the model, we can successfully classify a person face with or without mask.

VI. Contribution

All of us have equally contributed in completing the project.
1. Research:
Vikram:45 percent
Nikhil: 55 percent

2. Code:
Vikram:65 percent (Models and accuracy Plots)
Nikhil: 35 percent (Data processing and Analysis)
3. Report and Analysis:
Vikram:30 percent (Models and reason of performance)
Nikhil: 70 percent (Data analysis, plots visualisation and reason for performance)