The City College of New York

Topic: Face Mask Detection

CSC I6716 Computer Vision

Name: Jatin Jain
PhD Electrical Engineering
EMPLID # 24145250

TOPICS

1. INTRODUCTION

Object detection is a computer technology that is widely used with computer vision in various applications such as tracking movement of objects. Object can be a vehicle, a person, face detection or face recognition among many others. In this project, the object is face mask and implementation is of face mask detection.

In today's scenario we are amidst a global pandemic, COVID – 19. The spread of the disease has created an unrest throughout the world that has had a deep impact on the humanity. Given this, creation of safe work environment is of utmost importance and that involves wearing face masks and following social distancing according to Center for Disease Control and Prevention (CDC). The project is targeted towards the safety of the people by detection of face mask and to prevent the spread of the virus by monitoring in if people are actually following CDC guidelines.

The project includes building a good neural network which needs to detect from two sets of RGB images with masks and without masks and have better performance in terms of accuracy. The project also involves implementation of various Computer Vision algorithms and their effects on the accuracy of the model.

The Computer Vision algorithms can emphasize overall or the local characteristics of the image, clear the unclear image and emphasize on certain features of the image. It can improve image quality and enhance the contrast of the image.

Figure 1: Sample set of images

2. RELATED WORK

In recent years, object detection techniques using deep learning have potentially enhanced in handling complex tasks. But in an image understanding task, the convolutional neural network doesn't perform as in real world problems.

The Viola – Jones[1] object detection system can be used to detect any object but is especially common for facial detection. Zhu[2] also shared another widespread face detection algorithm which is neural network based detector.

The approach is not suited for our study, which requires in essentially to detect faces that have covered mouths with masks, scarfs etc. The approach is not useful in identifying different types of transparent masks that are recommended by CDC for deaf and hard of hearing people. Also, these approaches don't provide the accuracy of the model after implementation of the neural network. Several prototypes along with sensors using neural networks have been developed and are being used in New York City, but as discussed they don't provide the accuracy of the systems.

1. P. Viola and M. Jones, "Fast object detection using an enhanced cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I.

2. X. Zhu and D. Ramanan. Face detection, pose estimation,and landmark position in the wild. In IEEE CVPR, pages 2879–2886, 2012

3. APPROACH

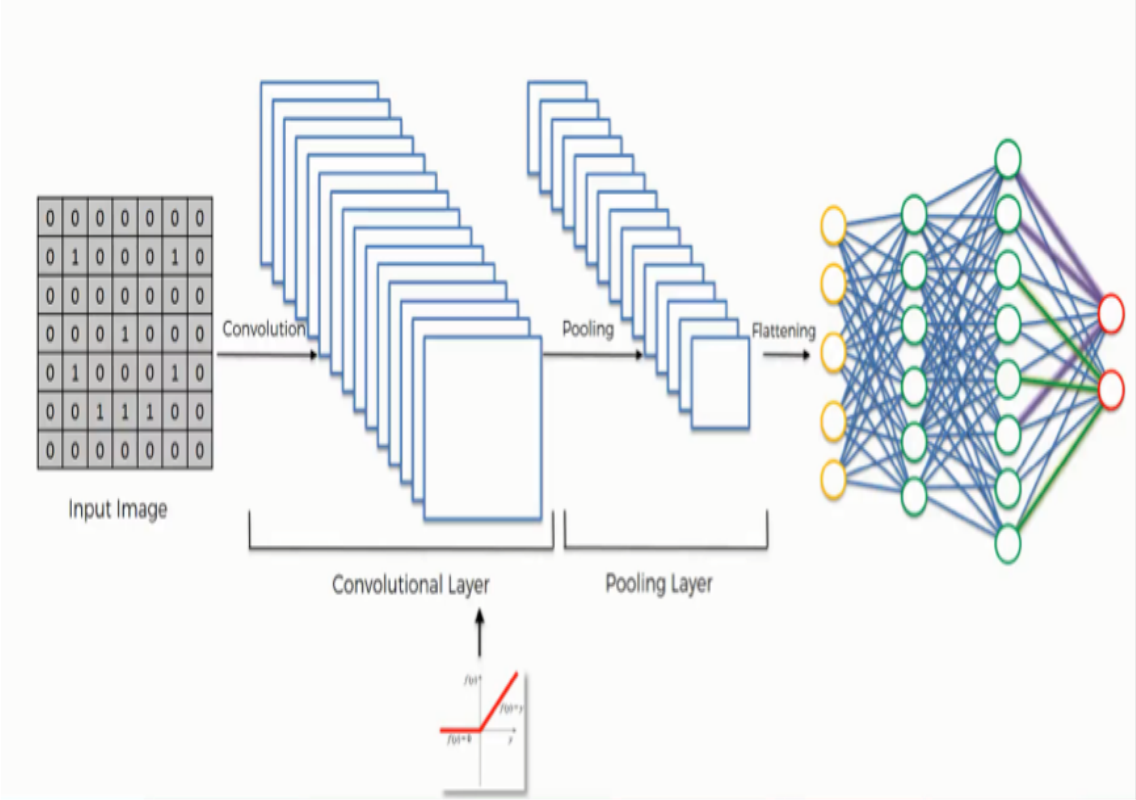
1. Implementation of CNN
2. Laplacian Operator
3. Sobel Operator

IMPLEMENTATION OF CONVOLUTION NEURAL NETWORK

- Data collection and pre-processing
  The model uses a dataset of RGB images consisting of images of different people, different images with different types of masks which are labelled and used for training of our model. The dataset used for training of our face mask detection model consists of 2000 images. The dataset is further divided into training and testing sets. The images in data collection are further labelled into Mask and No Mask.
- Experiment
  The model uses OpenCV and Tensorflow to train, for fast and efficient deep learning solution. The CNN takes an image as an input, processes it and classifies it. The model sees an image as an input array of pixels. The first layer is Convolution, which is used to extract features from an image.

  The second layer is Maxpooling which reduces the size/dimensions of an image without loosing it's important features.

  The third layer is flattening which transforms the 2D matrix of features into a vector that can be fed into a fully connected Neural Network classifier.



LAPLACIAN OPERATOR

Edge Detection is simply a case of trying to find the regions in an image where we have a sharp change in intensity or a sharp change in color, a high value indicates a steep change, and a low value indicates a shallow change.

Laplacian operator is a second derivative operator often used in edge detection. Compared with the first derivative-based edge detectors such as Sobel operator, the Laplacian operator may yield better results in edge localization. Unfortunately, the Laplacian operator is very sensitive to noise.

Definition of discrete Laplacian

$$\triangle f[m,n] =$$
$$f[m+1,n] + f[m-1,n] + f[m,n+1] + f[m,n-1] - 4f[m,n]$$

This operation can be carried out by 2D convolutional Kernel:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

SOBEL OPERATOR

A very common operator for doing this is a Sobel Operator, which is an approximation to a derivative of an image. It is separate in the y and x directions. At the x-direction, the gradient of an image in the x-direction is equal to this operator. The gradient for x-direction has minus numbers on the left-hand side and positive numbers on the right-hand side and we are preserving a little bit of the center pixels. Similarly, the gradient for y-direction has minus numbers on the bottom and positive numbers on top.

-1	0	+1
-2	0	+2
-1	0	+1

Gx

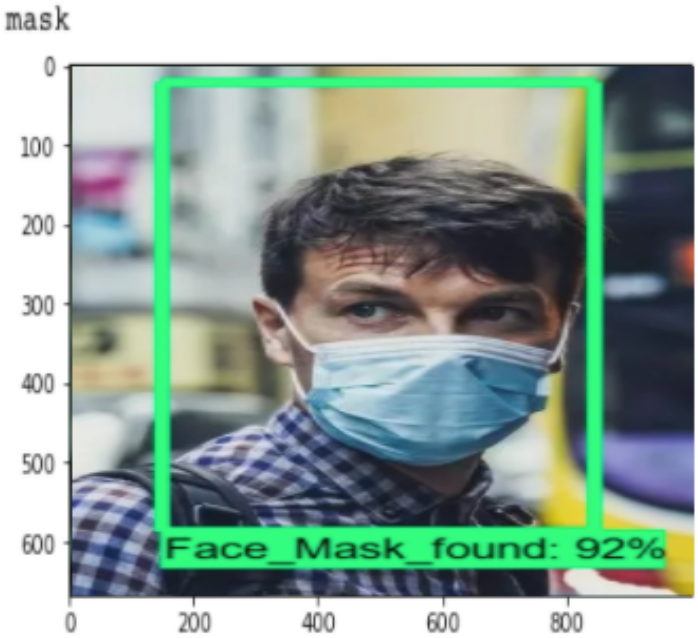
+1	+2	+1
0	0	0
-1	-2	-1

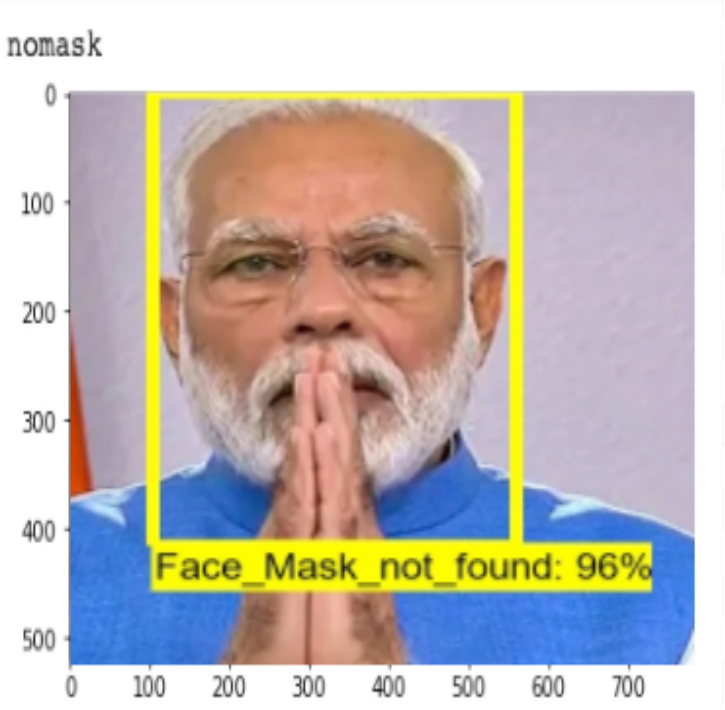
Gy

Edge Magnitude: $\sqrt{S_1^2 + S_2^2}$

4. IMPLEMENTATION AND ANALYSIS

1. Using Convolutional Neural Network Model: The model was able to identify images with mask and without mask.

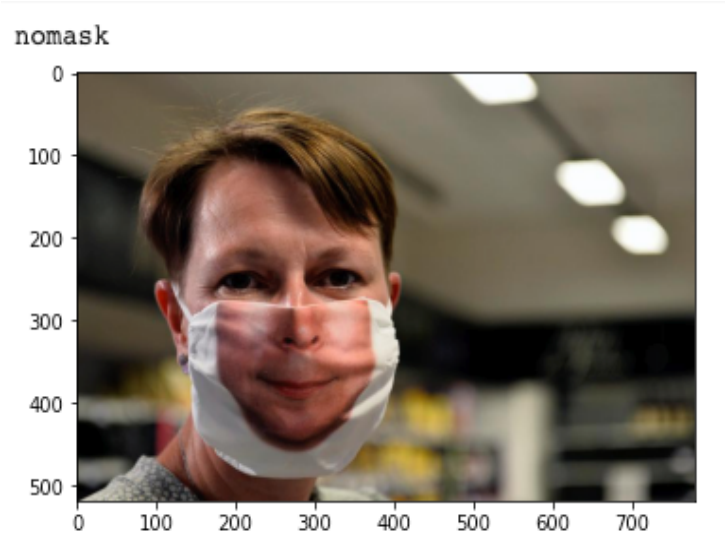


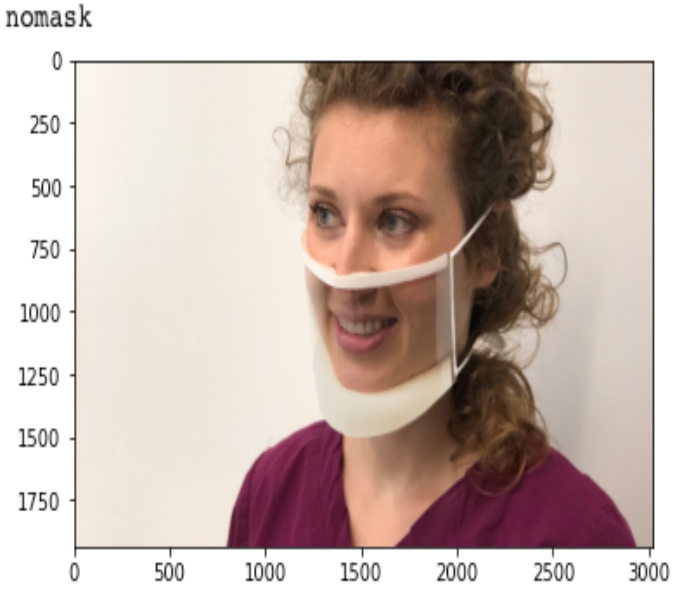


And the Accuracy of the model:

```
4/4 [==============================] - 1s 230ms/step - loss: 0.1750 - accuracy: 0.9200
test acc :0.9200000166893005 test loss:0.17502941191196442
```

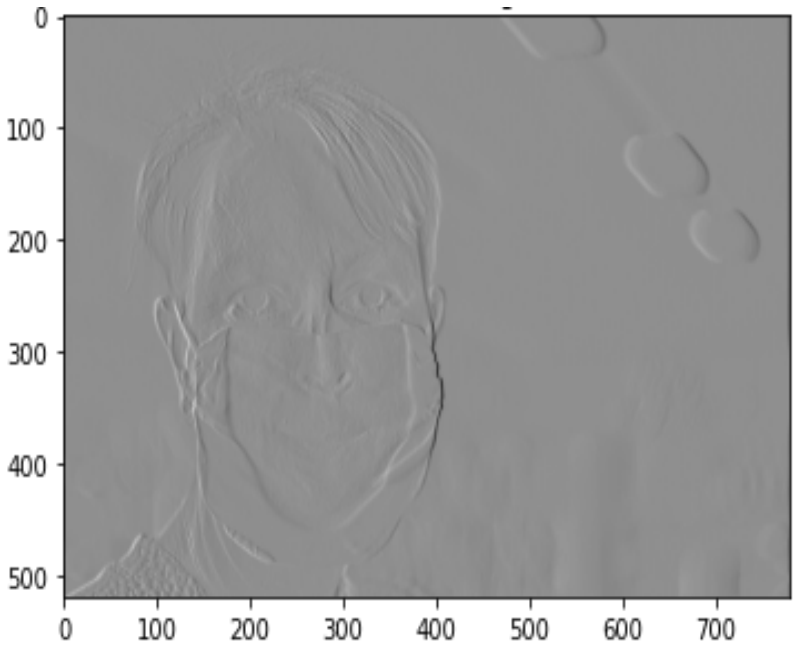
There were some exceptional cases for which model couldn't detect the mask:



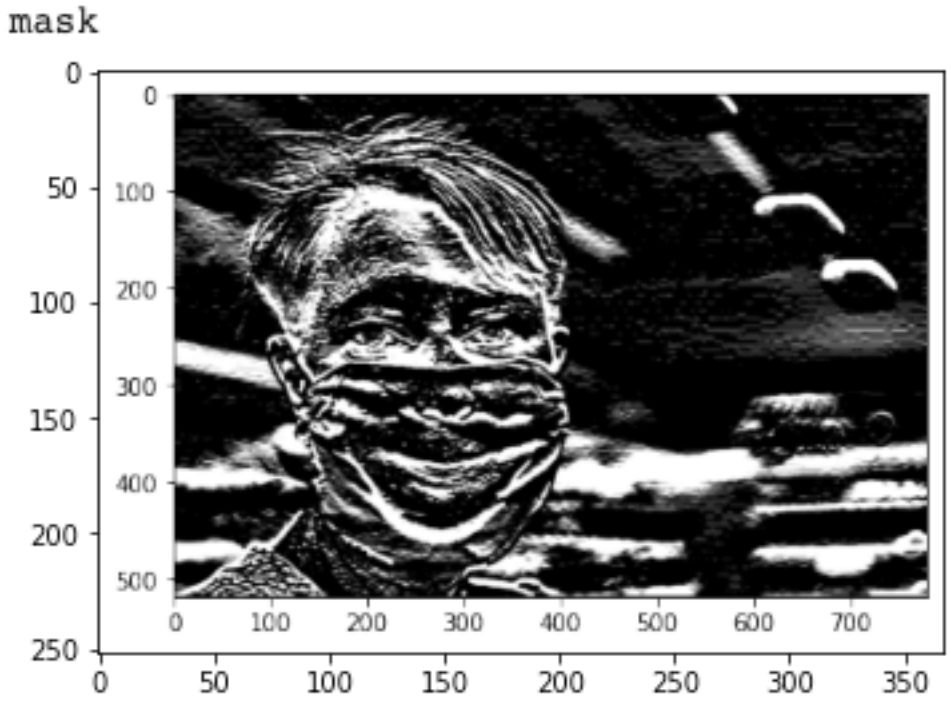


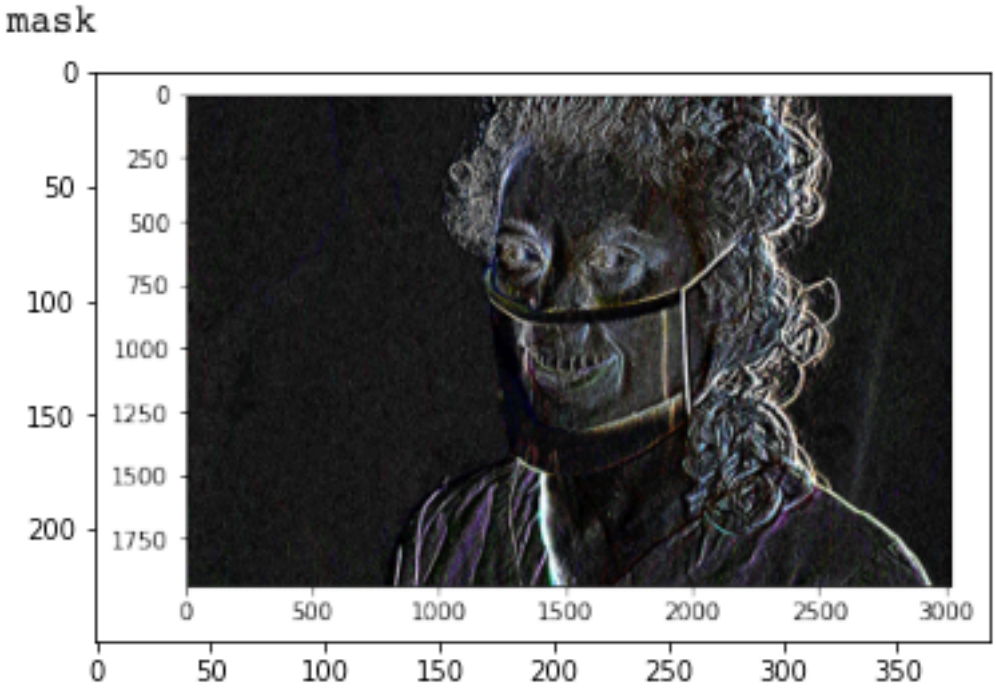
With Laplace Operator:

The model still couldn’t detect the face mask with Laplace operator:



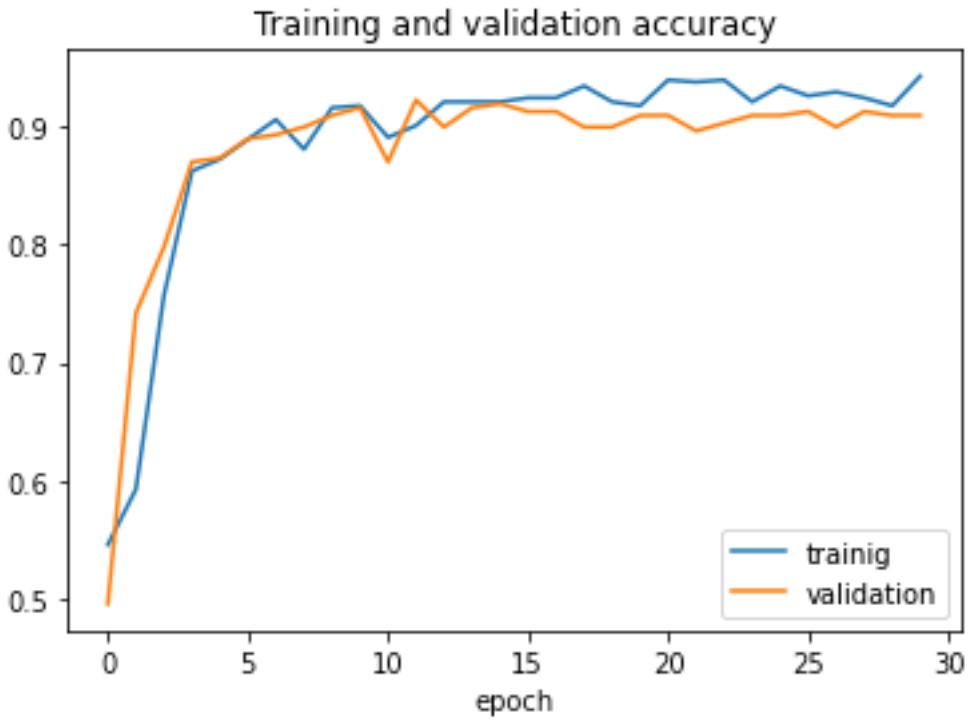
AFTER SOBEL OPERATOR:





And the accuracy of the model increased to:

```
4/4 [==============================] - 1s 221ms/step - loss: 0.1679 - accuracy: 0.9400
test acc :0.9399999976158142 test loss:0.1679234355688095
```



5. CONCLUSION

The project proposes an approach that uses Convolutional Neural network and Computer Vision algorithms to help create and maintain a secure working environment.

After working on the project, I could understand how the above application takes into consideration edge detection. How Sobel operator works and could understand the mathematical reasoning behind that.

There are several other cases of usage that can be included in this solution to offer a more detailed sense of safety. In future I intend to create cough and sneezing detection using deep learning with body gesture analysis.