Face Mask Detection using Open-Source Computer Vision Library and Scikit-Learn using Machine Learning

G. Poojitha[*1], S. Ganesh Kumar[2], G. Harsha Vardhan[3], M. Bhargavi[4], Mr. Vishwanadha Reddy[5]

[*1,2,3,4] Department of Computer Science & Engineering, NS Raju Institute of Technology, Visakhapatnam, Andhra Pradesh, India

[*5] Assistant Professor Department of Computer Science & Engineering, NS Raju Institute of Technology, Visakhapatnam, Andhra Pradesh, India

ABSTRACT

In this pandemic has rapidly affected our day to-day life disrupting the world trade and movements. Wearing a protective face mask has become anormal. In the near future, many public service providers will ask the customers to wear masks correctly to avail of their services. Therefore, face mask detection has become a crucial task to help global security. This project presents a simplified approach to achieve this purpose using some basic Machine Learning packages like OpenCV and Scikit-Learn. The proposed method detects the face from the image correctly and then identifies it has a mask on it or not. As a surveillance task performer, it can also detect a face along with the mask in motion. This method attains accuracy almost up to 90% on two data sets. We explore optimized values of parameters using the Viola Jones detection framework used to detect the presence of masks correctly without causing over-fitting.

Keywords: COVID-19, Coronavirus, Face mask detection, Voila-Jones algorithm, Open Computer Vision (OpenCV), Deep learning, Masked Face Detection Dataset (MFDD), Real-world Masked Face Recognition Dataset (RMFRD) and Simulated Masked Face Recognition Dataset (SMFRD).

I. INTRODUCTION

Coronavirus is a sort of sars-cov-2 infection that has stunned the world and was remembered for Indonesia in mid 2020, where this infection rose up out of China, Wuhan City. There are at present upwards of 65 nations tainted with the Covid, where consistently the quantity of cases in every nation keeps on expanding over the long haul. This Covid is possible gotten from creatures that are sent to people. Accordingly, the transmission cycle isn't known with conviction however it very well may be anticipated that transmission between people through drops and contact with infections delivered in drops. Thusly, everybody is expected to wear a veil or mouth cover so the gamble of contracting contact might be more

modest and this is likewise a work to forestall the spread of the infection. The utilization of veils is an administration work to spread the Covid and is likewise because of the disclosure of an immunization that actually demands investment.

In Indonesia, individuals are expected to wear covers, despite the fact that the sort of veil utilized is a cover produced using texture. In the utilization of veils, clients are supposed to be more cautious with the goal that the infection doesn't effectively enter. Convention rules are in many cases overlooked by people in general. To this end patients who are presented to the infection are not diminishing yet are expanding. This examination attempting to plan a face design recognition framework utilizing covers fully intent on having the option to help the gatherings worried in giving alerts to the public who have disregarded the convention rules by recognizing the pictures on whether the face having a veil. To plan this framework need a strategy that has a serious level of precision and quick processing is required.

In this study utilizing the Viola Jones strategy. Viola Jones strategy is a component choice technique that has a high identification rate for recognizing facial examples in pictures with low mistake rates. The cycle in this technique in face design discovery is by changing the information picture into another picture portrayal as a fundamental picture. The necessary picture comprises of various a lot of highlights. These elements will be chosen by AdaBoost to be utilized as a classifier part which will be utilized to group pictures. The ideal arrangement in the Viola Jones strategy is gotten by deciding the presence or nonattendance of certain haar highlights in a picture by choosing explicit haar highlights that will be utilized to change the edge, and the outpouring classifier as the last grouping decides the face region involving a veil in the picture of this technique. With the upsides of the strategy, a few investigations have created it as a method, for example, research directed

for the identification o on the face. Also, the exploration led for the presentation of human articles in the eyes, nose and mouth.

In this study the Viola Jones strategy can deliver a precision pace of 67.6%. The distinction of this exploration with past examinations is presently information investigation utilizing having the option to recognize blocked faces with covers, where the shapes are equivalent to the face forms in the format. The motivation behind applying this technique is to create an ideal framework configuration regarding exactness and registering. Testing this framework utilizes K-overlay cross-approval by noticing the degree of precision by adjusting the boundary values in the Viola Jones technique so the most elevated exactness is gotten for face pictures involving veils and low precision for face pictures without utilizing covers.

II. BACKGROUND

Anaconda-Navigator

Anaconda is an open source and free distribution of R and Python programming language for machine learning as well as data science projects. Therefore, it is known as a professional data science platform. It contains a powerful environment manager, which provides a different type of Python environment such as a Spyder, Jupyter notebook, and so on. Navigator can search for packages on Anaconda Cloud or in a local Anaconda Repository. It is available for Windows, macOS, and Linux.

Anaconda is a free and open-source distributor
Quickly download 7,500+ Python/R data science packages.
Manages a lot of Python libraries
It provides various environments by virtualization
Can easily deal with large data computing
It works properly without the need for any administrative privileges.
Visualize results with Matplotlib, Bokeh, Datashader , and Holoviews.

Open CV

OpenCV is the gigantic open-source library for the PC vision, AI, and picture handling and presently it assumes a significant part progressively activity which is vital in the present frameworks. By utilizing it, one can deal with pictures and recordings to distinguish articles, faces, or in any event, penmanship of a human. At the point when it coordinated with different libraries, for example, Numpy, python is equipped for handling the OpenCV exhibit structure for investigation. To recognize picture design and its different elements we use vector space and perform numerical procedure on these highlights.

The main OpenCV rendition was 1.0. OpenCV is delivered under a BSD permit and thus it's free for both intellectual and business use.

Sckit-learn

Sckit-learn is an open source Python library that executes a scope of AI,

Pre-handling, cross-approval and perception calculations utilizing a brought together point of interaction.

Significant highlights of scikit-learn:

·   Basic and proficient instruments for information mining and information examination. It highlights different grouping, relapse and bunching calculations including support vector machines, irregular timberlands, angle helping, k-implies, and so on.
·   Available to everyone and reusable in different settings.
·   Based on the highest point of NumPy, SciPy, and matplotlib.
·   Open source, monetarily usable - BSD permit.
·   Numpy

NumPy represents Numerical Python. NumPy is a python library utilized for working with clusters. It additionally has capacities for working in the area of direct polynomial math, Fourier change, and lattices. NumPy was made in 2005 by Travis Oliphant. It is an open-source venture, and you can utilize it uninhibitedly. NumPy plans to give an exhibit object that depends on 50x quicker than conventional Python records. It is enhanced to work with the most recent CPU structures. NumPy clusters are put away at one nonstop spot in memory not at all like records, so cycles can get to and control them effectively.

III.   PROBLEM DEFINITION

1. From that point, we will audit the dataset we will use to prepare our custom facial covering locator.
2. I will then, at that point, tell you the best way to execute a python content to prepare a facial covering finder on our dataset by carrying out viola-jones calculation.
3. We will utilize this python content to prepare a facial covering locator and survey the outcomes.
4. Given the prepared COVID-19 facial covering locator facial covering indicator, we will continue to execute two additional extra python scripts used to:Detect COVID-19 face mask in images Detect face masks in real-time

We will wrap up the post by looking at the results of applying our face mask detector.

A. Face identification calculation

One of the well known calculations that utilization a component based approach is the Viola-Jones calculation and here I am momentarily going to examine Face discovery utilizing Viola-Jones.

B. Viola-Jones calculation is named after two PC vision specialists who proposed the strategy in 2001, Paul Viola and Michael Jones in their paper, "Fast Object Detection utilizing a Boosted Cascade of Simple Features". Notwithstanding being an obsolete

C. Framework, Viola-Jones is very strong, and its application has demonstrated to be especially outstanding progressively face discovery. This calculation is agonizing to prepare however can distinguish faces progressively with amazing rate.

D. Given a picture (this calculation deals with grayscale picture), the calculation takes a gander at numerous more modest sub districts and attempts to track down a face by searching for explicit elements in each sub locale. It requirements to check a wide range of positions and scales on the grounds that a picture can contain many countenances of different sizes. Viola and Jones utilized Haar-like highlights to recognize faces in this calculation.

The Viola Jones calculation has four fundamental stages, which we will talk about in the segments to follow:

E. Selecting Haar-like elements

F. Creating a basic picture

G. Running AdaBoost preparing

Head Component Analysis (PCA):

PCA is a measurable methodology that utilizes a symmetrical change that switches a bunch of connected factors over completely to a bunch of uncorrelated factors. PCA is the most generally involved device in exploratory information examination and in AI for prescient models. Besides, PCA is an unaided factual strategy used to inspect the interrelations among a bunch of factors. It is otherwise called a general variable examination where relapse decides a line of best fit.

Uses of PCA:

· It is utilized to find between connection between factors in the information.

· It is utilized to decipher and envision information.

· The quantity of factors is diminishing it simplifies further investigation.

· Envisioning hereditary distance and relatedness between populations is frequently utilized

Goals of PCA:

· It is essentially a non-subordinate method in which it decreases trait space from an enormous number of factors to fewer elements.

· PCA is fundamentally an aspect decrease process yet there is no assurance that the aspect is interpretable.

· The fundamental errand in this PCA is to choose a subset of factors from a bigger set, in view of which unique factors have the most elevated relationship with the chief sum.

Head Axis Method:

PCA fundamentally look through a straight blend of factors with the goal that we can separate most extreme difference from the factors. When this interaction finishes it eliminates it and looks for one more direct blend that gives a clarification about the most extreme extent of outstanding difference which essentially prompts symmetrical elements. In this technique, we examine complete fluctuation.

Eigenvector:

A non-zero vector stays lined up after framework increase. We should assume x is an eigenvector of aspect r of framework M with aspect r*r in the event that Mx and x are equal. Then we want to tackle Mx=Ax where both x and An are obscure to get eigenvector and eigenvalues.

Under Eigen-Vectors we can say that Principal parts show both normal and interesting fluctuation of the variable. Essentially, it is fluctuation centered approach trying to imitate absolute difference and relationship with all parts. The central parts are fundamentally the direct mixes of the first factors weighted by their commitment to make sense of the difference in a specific symmetrical aspect.

Eigen Values:

It is fundamentally known as trademark roots. It essentially gauges the change in all factors which is represented by that element. The proportion of eigenvalues is the proportion of informative significance of the elements as for the factors. On the off chance that the component is low, it is offering less to the clarification of factors. In straightforward words, it estimates how much change in the complete given data set accounted by the component. We can compute the element's eigenvalue as the amount of its squared component stacking for every one of the factors.

A. Support Vector Machine:

Support Vector Machine or SVM is one of the most famous Supervised Learning calculations, which is utilized for Classification as well as Regression issues. In any case, basically, it is utilized for Classification issues in Machine Learning.

The objective of the SVM calculation is to make the best line or choice limit that can isolate n-layered space into classes so we can undoubtedly put the new data of interest in the right classification later on. This best choice limit is known as a hyperplane.

SVM picks the outrageous focuses/vectors that assistance in making the hyperplane. These outrageous cases are called as help vectors, and thus calculation is named as Support Vector Machine. Consider the beneath graph in which there are two unique classifications that are ordered utilizing a choice limit or hyper SVM can be perceived with the model that we have utilized in the KNN classifier. Assume we see an unusual feline that likewise has a few highlights of canines, so assuming we need a model that can precisely recognize whether it is a feline or canine, so such a model can be made by utilizing the SVM calculation. We will initially prepare our model with bunches of pictures of felines

and canines so it can find out about various highlights of felines and canines, and afterward we test it with this peculiar animal. So as help vector makes a choice limit between these two information (feline and canine) and pick outrageous cases (support vectors), it will see the outrageous instance of feline and canine. Based on the

Support vectors, it will order it as a feline.

SVM can be of two sorts:

Linear SVM: Linear SVM is utilized for directly distinct information, and that implies if a dataset can be grouped into two classes by utilizing a solitary straight line, then, at that point, such data is named as straightforwardly recognizable data, and classifier is used called as Linear SVM classifier. Non-direct SVM: Non-Linear SVM is utilized for non-directly isolated information, and that implies on the off chance that a dataset can't be characterized by utilizing a straight line, such information is named as non-straight information and classifier utilized is called as Non-linear SVM classifier.

Hyperplane and Support Vectors in the SVM algorithm:

Hyperplane: There can be various lines/choice limits to isolate the classes in n-layered space, yet we want to figure out the best choice limit that assists with ordering the data of interest. This best limit is known as the hyperplane of SVM.

The elements of the hyperplane rely upon the highlights present in the dataset, and that implies in the event that there are 2 elements (as displayed in picture), hyperplane will be a straight line. Furthermore, in the event that there are 3 highlights, hyperplane will be a 2-aspect plane. We generally make a hyperplane that has a most extreme edge, and that implies the greatest distance between the data of interest.

Support Vectors: The important pieces of information or vectors that are the nearest to the hyperplane and which influence the place of the hyperplane are

named as Support Vector. Since these vectors support the hyperplane, thus called a Support vector.

Linear Svm:

The working of the SVM algorithm can be understood by using an example. Suppose we have a dataset that has two tags (green and blue), and the dataset has two features x1 and x2. We want a classifier that can classify the pair(x1, x2) of coordinates in either green or blue. Consider the below image:
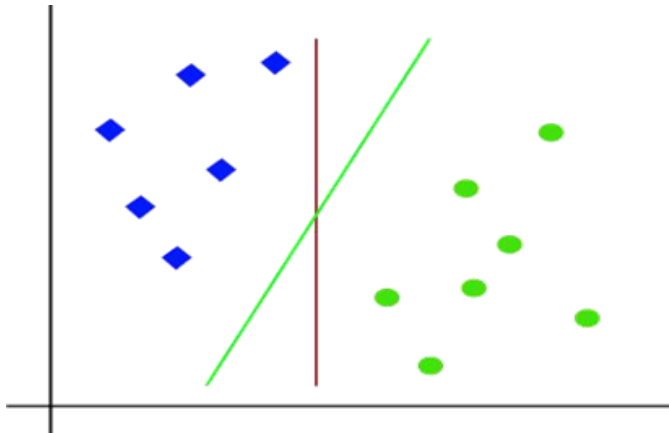
So as it is 2-d space so by just using a straight line, we can easily separate these two classes. But there can be multiple lines that can separate these classes. Consider the below image:
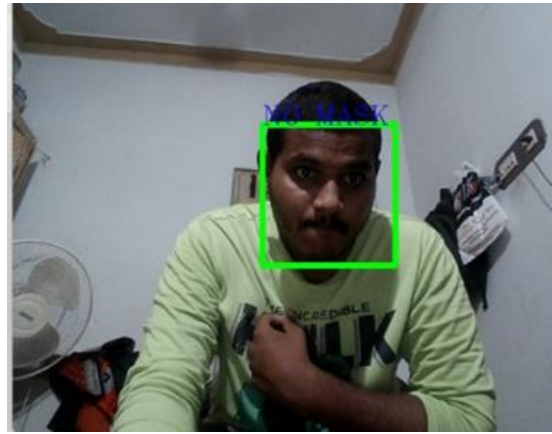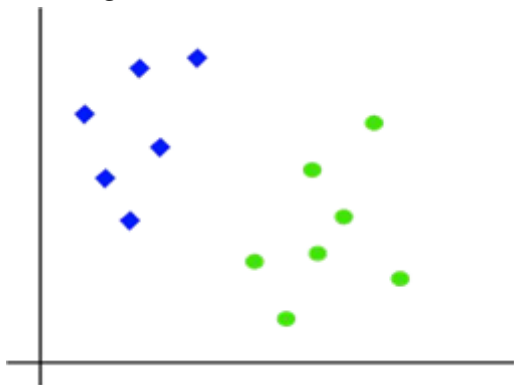
IV. CONCLUSION

With the rising number of COVID cases from one side of the planet to the other, a framework to supplant people to check veils on the essences of individuals is extraordinarily required. This application fulfills that need. This application can be utilized openly puts at passage level in rail route stations, shopping centers, emergency clinics, and so forth. It will be of an extraordinary assistance in organizations and immense foundations where there will be a ton of laborers. This application will be an incredible assistance there since it is not difficult to recognize faces and shows that he/she has the cover. So this is exceptionally helpful to forestall the Coronavirus infection by making mindfulness by putting cover.

V. REFERENCES

[1]. https://www.mygreatlearning.com/blog/real-time-face-detection/

[2]. https://www.javatpoint.com/machine-learning-support-vector-Machine-algorithm

[3]. Reshaping Inputs for Convolutional Neural Networks -Some common and uncommon methods DOI: 10.1016/j.patcog.2019.04.009

[4]. Big Data Concepts and Techniques in Data Processing DOI: 10.26438/ijcse/v6i10.712714

[5]. Convolutional neural networks: an overview and application in radiology DOI: 10.1007/s13244-018-0639-9

[6]. Amalgam Approach to Computer Vision - Deep

[7]. Learning and IoT Approaches for Aerial View Cultivating Monitoring – IEEE- COMSOC- MMTC Communications- Frontiers

Cite this article as :