Indian Institute of Information Technology, Allahabad

Department of Information Technology

PROJECT REPORT

Attendance system using real-time face recognition

Submitted By:

Harsh Kumar(IIT2017119)
Krishna Jaiswal(IIT2017117)
Anshul Singh(IIT2017119)

Under the Supervision of

Dr. Vrijendra Singh

HOD(IT)

Declaration By The Candidates

We hereby declare that the work presented in this semester project report of B.Tech (IT) 6th Semester entitled “ATTENDANCE SYSTEM USING REAL TIME FACE RECOGNITION", submitted by us at Indian Institute of Information Technology, Allahabad, is an authenticated record of our original work carried out in February 2020 under the guidance of Dr. Vrijendra Singh.

Due acknowledgements have been made in the text to all other material used. The project was done in full compliance with the requirements and constraints of the prescribed curriculum.

Date :
Place : IIIT Allahabad

Supervisor:
Dr. Vrijendra Singh
HOD(IT)

Abstract

This report describes the 6th semester project our group has worked on, titled "Attendance System using Real-time Face Recognition". The study aims to mark attendance without much user effort. The first step that is required is face detection, i.e., detecting faces in an image, video or real time coverage consisting of different types of objects (distinguishing faces from non-faces) .

Once the face is detected, feature extraction, i.e, determining the uniqueness of the face by taking out the features, is performed. The Face Extraction is performed using OpenCV library in python. Once the extracted face is obtained a neural network is used for identification of the extracted image.

Contents

1. Introduction

A facial recognition system is a technology which is used for the identification and verification of humans from a digital photo or video frame by characteristics of their faces. Human face recognition is the fastest biometric technology which is widely used in our day to day applications, such as security systems, video monitoring systems, gaming etc. This project aims at providing an ease and efficient system integrated with face detection to record the attendance of the students.

A naive approach for the above will be to arrange the face detector at a certain place and the student needs to get his face recognised by coming towards it and thereby getting marked as present but this method involves limitations and is time consuming.

So in order to overcome this, we've decided to mark the attendance of the student automatically by fixing the camera in such a way that it can capture the entire class and mark the attendance.

2. Motivation

We intend to learn new technologies and apply them to our daily lives to make them easier. Machine learning has had a very huge impact on human lives. The use of machine learning algorithms to extract the features of an image is already in the market for quite some time now
but there are not many applications in the market which leverage this technology. Often when the attendance is marked, it consumes a lot of time and human effort and eventually comes out to be with many false attendances("proxies"). We aim to solve this problem of manually taking attendance and provide a model which would resolve the difficulties occurred in the process of taking attendance.

3. Problem Statement

Attendance Management System using Face Recognition is a modern enhancement in software development for student attendance in various educational institutions which overcomes the limitation of the traditional pen-paper attendance management system by creating an interface that can be used to mark attendance of the students automatically to save time and reduce human effort.

4. Comparison With Other Models

There are a few pre-existing models based on Attendance Management using face recognition but most of them include arranging a face detector at a certain place and the students need to come towards the detector and the attendance will be marked as soon as the face is recognized. This involves various limitations and crunches a lot of time.

So in order to overcome this, we've not only decided to record the attendance of students in classroom automatically, but also to integrate the system with a control center for the faculty to access the student information and to monitor the working duration of the system based on
the class timings.

This proposed system aims at providing an ease and efficient system integrated with face detection followed by face recognition to record the attendance of the students.

5. Review of Literature Survey

With help of the literature survey done for face recognition, we realized that the basic steps involve face detection and face recognition. During the face detection process the captured image is classified into regions like “face” and “non face”.

1. In [1] the authors have evaluated Adaboost classifier with Haar, LBP and SVM classifier with HOG face detection methods. The experiments conducted on various datasets shows that the mean accuracy of Haar is highest.

2. In [2], the face detection approaches like are holistic approach and feature based approaches were considered. In the Holistic approach, the entire face region is considered as input to the face detection system. In feature based approach, the features of face such as nose and eyes are then taken as input in face detection systems.

3. In [3], authors explained MTCNN which is enhancement of convolutional neural networks to detect and align the images.

4. In [4] ,authors explained how Histograms of Oriented Gradients are useful for Human Detection . Viola and Jones has a good face detector for most of the applications where the face is in a controlled environment (facing the camera frontally, good illumination conditions, etc.).But whereas HoG features are capable of capturing the pedestrian or object outline/shape better than Haar features.

For face recognition, we need to use different algorithms, like PCA and Eigenface techniques which use the concept of region of interest ,machines techniques.

1. In [5] the authors have proposed the use of Principal Component Analysis. PCA is a technique used for dimensionality reduction and thus reducing the training time of the dataset and even can efficiently extract the features of the face . The eigen vectors generated by training the model on facial space is thus reduced into a lower subspace . Various experiments were conducted and the found accuracy is 96.38%.

2. In [6] the proposal put forward by the authors is that the model learns an embedding system based on the Euclidean system. We must train the network the squared matrix of the embedded space to maximize the distance between the eigenvectors of two distinct faces and also classify between large and small faces i.e. distance between their facial features should be proportional.

Based on the above literature survey we concluded that implementing the MTCNN for face detection would be apt and when it comes to face

recognition the results show that implementing FaceNet will output satisfactory results.

6.Technologies Used and their Definitions

6.1.Face Detection and Recognition:

Face detection is the technology we used in our project to separate out and crop the human faces from the supplied images.Using the distinct features of the face i.e. the nasal part and distances from the eye and various other features it detects the facial part of the body.

Face recognition system is the technology to differentiate the set of images and mark its distinct features . From an ongoing video a frame is selected and the machine crops and matches the features of the face with the data set of eigenvectors and returns whether a match is found or the person is not present in the dataset.

6.2 Convolution Neural Network :

Convolution Neural Network also known as Convo Nets or associated with the abbreviation CNN.This is a very useful part of the face recognition technology. Firstly face detection and secondly can classify between a set of given facial features.Mainly used for image , video recognition , image classification .The neurons present in the network work collectively as division of three layers which are classified on the dimensions namely height,width,depth.The end product generated from CNN architecture is that for a given image it will store the features of that face in the form of an eigenvector.

6.2.1 Layers in CNN:

1. Convolution Layer : It is the starting layer of the Convo Net which is used for scraping out the features from a given dataset image. It keeps learning and developing itself and also maintains a concrete relation between eigenvectors which are the representation of the facial features and input data squares (which are stores are from a matrix format).. In this layer we take a small window size [typically of length 5*5] that extends to the depth of the

input matrix. The layer consists of learnable filters of window size. During every iteration we slid the window by stride size [typically 1], and computed the dot product of filter entries and input values at a given position. As we continue this process we will create a 2-Dimensional activation matrix that gives the response of that matrix at every spatial position.

2. Pooling Layer : We use a pooling layer to decrease the size of the activation matrix and ultimately reduce the learnable parameters.
There are two types of pooling :
a) **Max Pooling** : In max pooling we take a window size [for example window of size 2*2], and only take the maximum of 4 values. We will slide this window and continue this process, so we will finally
get an activation matrix half of its original size.
b) **Average Pooling** : In average pooling we take the average of all values in a window.

3. Fully Connected Layer : In convolution layer neurons are connected only to a local region, while in a fully connected region, we connect all the inputs to neurons.

4. Final Output Layer : After getting values from the fully connected layer, we connect them to the final layer of neurons[having count equal to total number of classes],that will predict the probability of each image to be in different classes.

6.3 Face Detection :

6.3.1.Ada Boost Approach:

It is short for Adaptive Boosting.It is used alongside with many algorithms to improve performance. It works by choosing a base algorithm and improving algorithms by iteratively checking for incorrectly predicted values in the training set.

6.3.2.Haar Cascade:

It is proposed by Paul Viola and Michael Jones in their paper, "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001.The training is

done by superimposing the set of positive images over the set of negative images.

6.3.3 Face detector from dlib:

● Dlib is a toolkit for making real world machine learning and data analysis applications in C++.
● One of the model in dlib is widely used face detection model, based on HoG features and SVM.The model is built out of 5 HOG filters – front looking, left looking, right looking, front looking but rotated left, and a front looking but rotated right.But it does not work for side face and extreme non-frontal faces, like looking down or up but works very well for frontal and slightly non-frontal faces.
● Another model present in dlib is based on Maximum-Margin Object Detector (MMOD) with CNN based features. It works fine for different face orientations and is robust to occlusion but is very slow on CPU and very fast on GPU

6.4 Face Recognition:

6.4.1.Principal Component Analysis:

Principal component analysis is a procedure by which the the dimensionality of an informational collection containing many correlated variables,vigorously or delicately, is diminished while holding the variety present in the information set.Thus the variables present are reduces to another set of new variables that are orthogonal to each other.These new set of variables are called Principal Components. These principal components are the eigenvectors of a covariance matrix, and hence they are orthogonal . Overall PCA helps us to discover the dimensions of the data set which actually matters or contains most of the information.

6.4.2 Eigenfaces:

Eigenfaces are a set of eigenvectors which are created when a facial recognition model runs on a set of images.These can be basically added to mean eigen vector to create new facial features.
Mathematically,

$$F = F_m + \sum_{i=1}^{n} \alpha_i F_i$$

where,
F is the face that is newly generated,
F_m mean face vector we will use,
F_i ith eigen vector from the set of eigenfaces,
α_i the weight of each feature we want to keep in the newly generated eigenface.These can be -ive or +ive.

6.4.3 OpenFace :
OpenFace is a python implementation of Face Recognition with deep neural networks and is based on Facenet which trains its output be a compact 128-D embedding using a triplet based loss function.FaceNet uses a deep convolutional network.

6.5 OpenCV :

OpenCV(Open Source Computer Vision) is an open source library of programming functions used for real-time computer-vision. It is mainly used for image and video processing and analysis for features like face and object recognition. It is written in C++ which is its primary interface, however bindings are available for Python, Java, MATLAB/OCTAVE.

6.6 Triplet Loss:

L - representation of Triplet Loss.
For understanding this concept we need to know the following :-

Negative image - This is the complement of the given image i.e. the eigenthe vector has been inverted . The 0's become 1 and the 1's become 0.

Positive image - Original image

Anchor point - It defines the positioning of the pixel with respect to the Kernel.
Triplet loss is to be minimised . This happens when the distance between anchor image and negative image in embedding space is greater than the distance between anchor image and positive image by at least by some given minimum(N),

$$\sum_{i=1}^{N} \left[(f_i^a - f_i^p)^2 - (f_i^a - f_i^n)^2 + N \right]$$

f^p - embedding of the positive image

f^n - embedding of the negative image

f^a - embedding of the anchor image

Now if we create the triplets which can satisfy the given relation would result in taking the model farther from convergence and result in greater complexity . We need to select better triplets that would result in quicker convergence and thus imporving the accuracy of the model.

7. Methodology

7.1 Architecture:

● The working of this Attendance Management System is very simple and user friendly. The hardware requirements includes a high definition camera with the field of view sufficient enough to cover the classroom and supported

by openCV.This camera has to be fixed in the classroom at a suitable position such that it can cover the required area.

● The pictures captured by the camera are enhanced for further processing.

● During this process of the enhancement, the picture is first transformed in grayscale image, and then respective Face Recognition algorithm is applied.

● Then after detection of faces, each student's face will be cropped from that image and will be given for recognition. Thus the features extracted are compared to that of the database to mark the attendance.

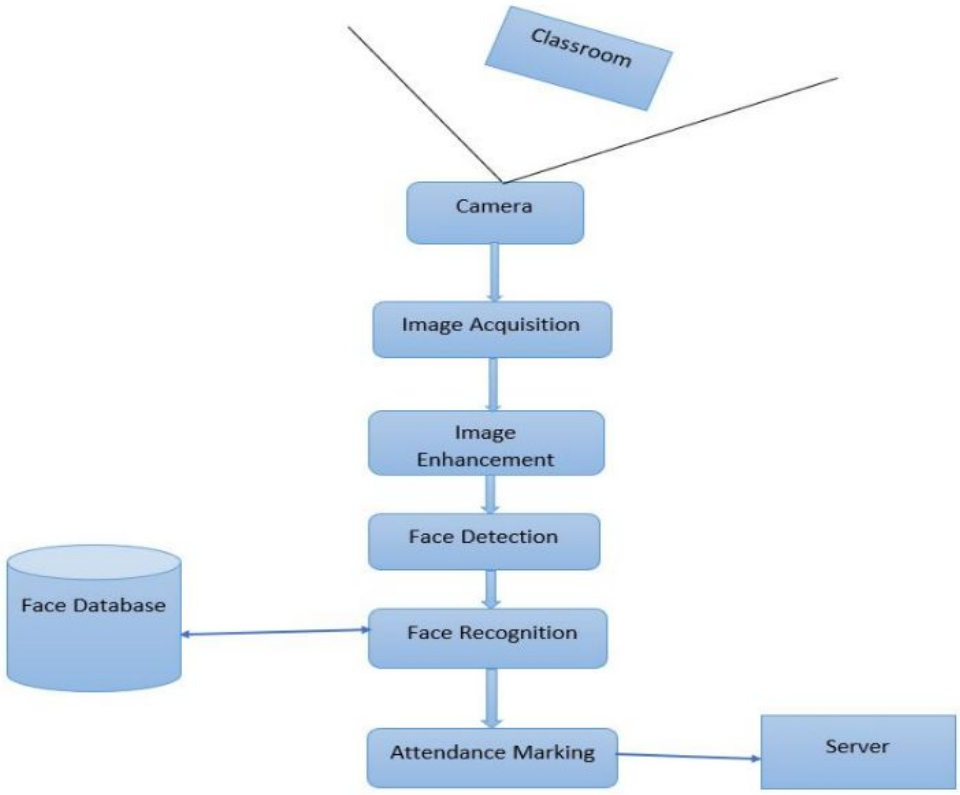


Fig. 1: Architecture of face recognition based attendance system.

7.2 Methodology:

For implementing the automated face recognition system, we need to follow some particular methodologies. The certain steps need to be performed for this process.

1. Extracting the embeddings for the collected images of the candidates.
2. Detect faces from the data collected from the video surveillance
3. Recognise the detected faces and mark attendance

7.2.1 .Extracting the embeddings

1. The image is passed through the detector of dlib to detect the face.
2. The detected face is made to undergo the face alignment processes.
This preprocessing step is very important for the performance of the neural network.
3. We then extract 128-dimensional representations, or embeddings, of faces by using the face detection module of dlib which in turn uses HOG for face extraction.
4. These embeddings serve as our database.

7.2.2 Recognise the detected faces

1. The obtained data is compared against the data in the database.
2. The distance between them is calculated and it is shown that the distance between the two images of the same person is smaller than the distance between images of different candidates.
3. To find the optimal value for τ (variable related to the distance vector), the face verification performance must be evaluated on a range of distance threshold values. At a given threshold, all possible embedding vector pairs are classified as either the same identity or different identity and compared to the ground truth.Thus the final τ value is where the maximum accuracy occurs.
4. Given the distance threshold τ , face recognition is now done by calculating the distances between an input embedding vector and all embedding vectors in a database.
5. The input is assigned the label (i.e. identity) of the database entry with the smallest distance if it is less than τ or label unknown otherwise.
6. Thus the attendance is marked with the help of the assigned label to the input image and sheets are edited.

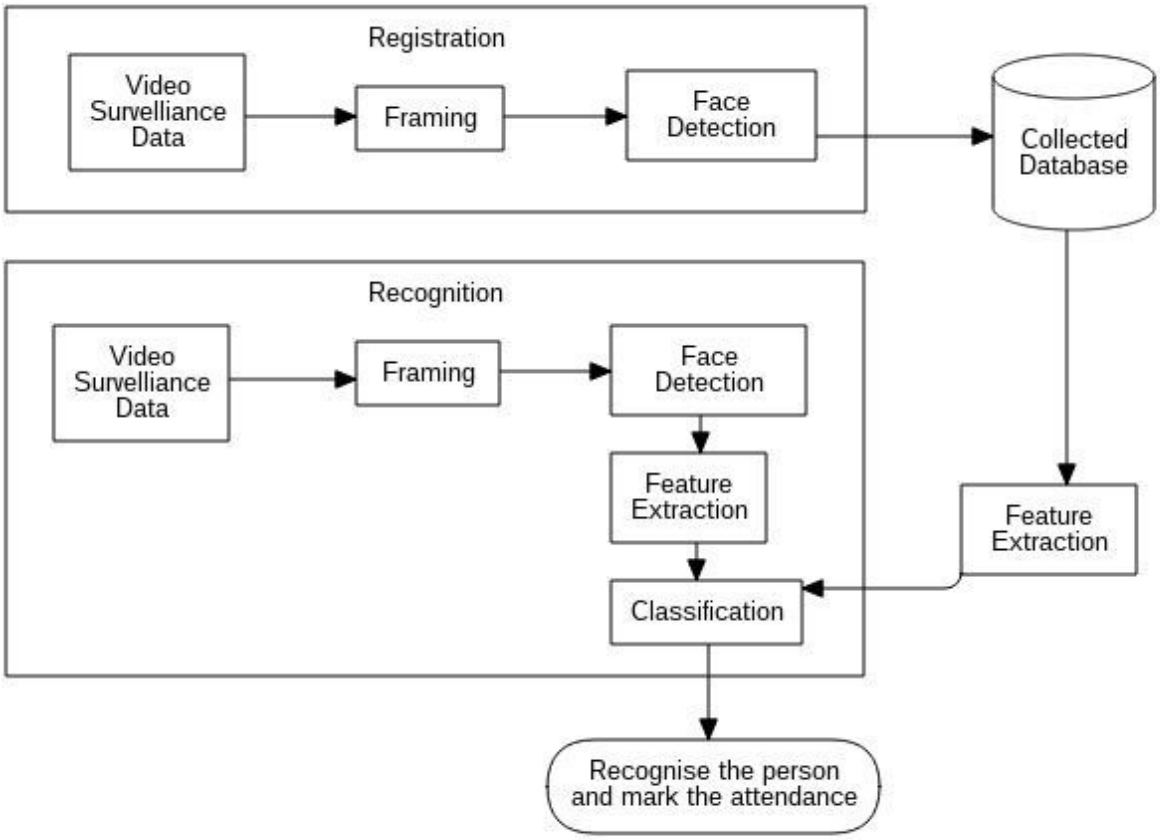
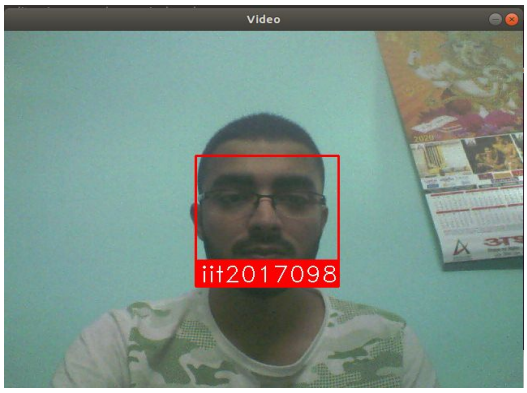


Fig 2 : The flow chart of proposed system

8.Implementation Plan

1. Acquainting ourselves with facial detection and recognition algorithms and capturing images for dataset generation.
2. We used face detection utilities to detect faces.
3. We worked on and studied face recognition models to recognize faces using the stored data .
4. We trained our model on various images and it learned essential features for classification .
5. After getting proper accuracy and good results we integrated our backend with a UI for the student side and the command line interface for the Admin side.

9. Results and Discussion

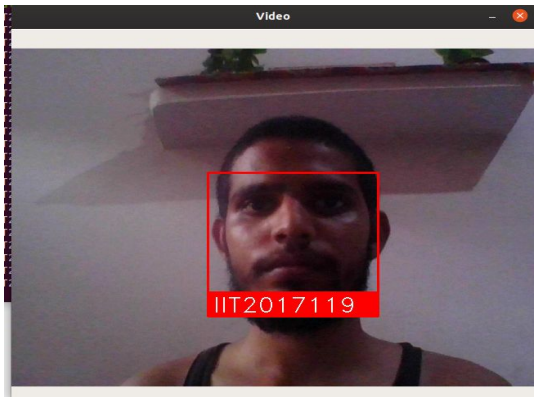


(output screenshot after recognition)

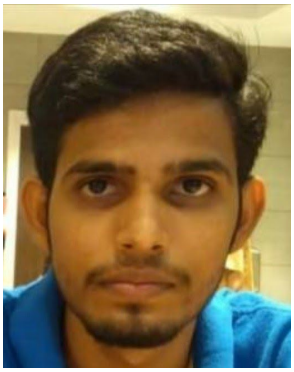
(image taken in march 2019)

● The image taken in march 2019 was used for training the model. It's clear that the student had a different hair style and was wearing no eye glasses in the old photo. The model has still given accurate results despite the changes in the face.

Below is another test case, with differences in facial hair.

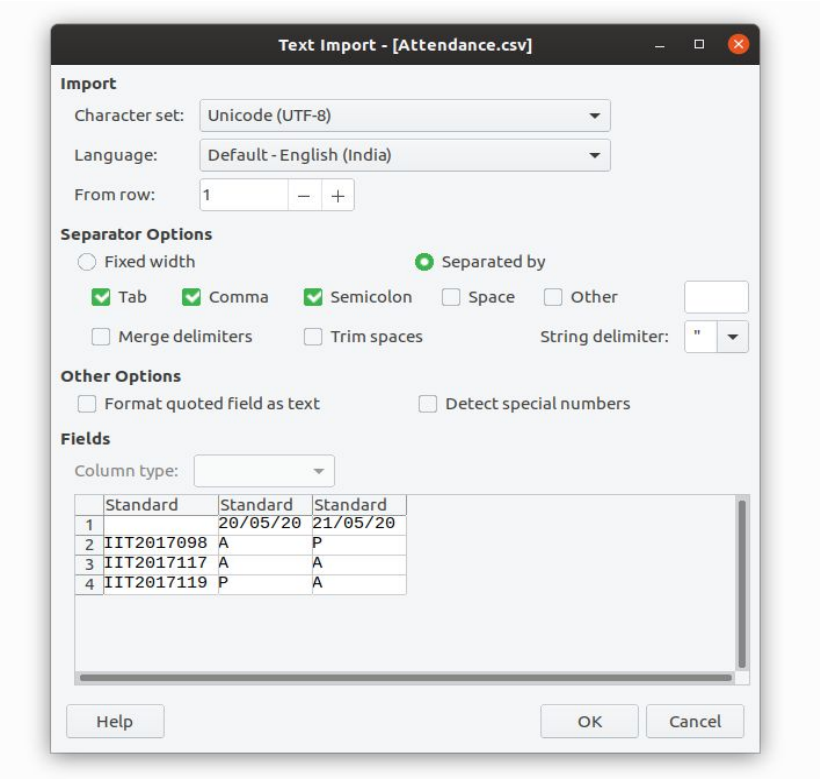


(output screenshot after recognition)



(old image taken before)

● The proposed system is able to reduce the hectic work of manually taking attendance thus increasing actual lecture time of the Professor.
● The system should also follow the attendance criteria proposed by the Institute.



- Only the same attendance sheet is updated whenever the model is used. ‘A’ is for absent and ‘P’ is for present.

10. Conclusion and Future Scope

● The system takes pictures and generates attendance with an accuracy of 96.38% and it fluctuates mildy with the given dataset.
● This system can be further extended into installing cameras into each of the classes and thus marking attendance on demand.
● This system can be further implemented in the industrial sector to keep a check on the employees of the company.
● It can also be used for security systems for houses i.e. if the model recognises the person then the door automatically opens for the family members living in the house.

11.References

1. Faizan Ahmad , Aaima Najam and Zeeshan Ahmed : Image-based Face Detection and Recognition: “State of the Art”- IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 1, November 2012.

2. K. Susheel Kumar, Shitala Prasad, Vijay Bhaskar Semwal, R. C. Tripathi : Real Time Face Recognition using AdaBoost Improved Fast PCA Algorithm : IJAIA, vol. 2, no. 3, July 2011.

3. Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Senior Member, IEEE, and Yu Qiao,Senior Member,IEEE :Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks

4. Navneet Dalal and Bill Triggs:Histograms of Oriented Gradients for Human Detection

5. A. M. Patil, Satish R. Kolhe, Pradeep M. Patil:"Face Recognition by PCA Technology”: Second International Conference on Emerging Trends in Engineering & Technology (ICETET) .

6 . Florian Schrofff,Dmitry Kalenicheko,James Philbin: FaceNet:A Unified Embedding for FAce Recognition and Clustering

7.https://docs.opencv.org/3.1.0/d7/d8b/tutorial_py_face_detection.html

8. OpenCV : https://www.learnopencv.com/