

Class Monitoring System Tools MTCNN and Haarcascade Classifier

Aditya Vikram Bhattacharya¹, Mrinalini Khanna², Akshay Tripathi³, S.Murugaveni⁴

^{1,2,3,4}Dept. of Information and Telecommunication, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Kancheepuram District, Tamil Nadu, India

*Corresponding Author Email: ¹adityavikrambhattacharya91@gmail.com

Abstract

The project aims towards the assistance of teachers at the time of taking attendance. The system solely focuses on face detection and recognition. The tools used to device the system are API's offered by Python 3.6, Open CV (for detection) and a few cognitive tools provided by Azure. The basic idea behind the project is face recognition linked to a database backend. The information of the student attending the class is stored here. The entire attendance is associated with two types of time stamps incorporated at the server end. The time stamp helps to keep a track of the hour conducted and the time for which number of people attended the class. Exceptions in the time stamp would be incorporated in order to cater for the students leaving the class or trying to bunk the class. In case of further exceptions in the time stamp will be scope of further development of the system. All queries or conditions of the students will be answered by the system on communication with the admin. If the admin finds that the system was at fault then it can always be fixed by the admin for smooth functioning of the class monitoring system.

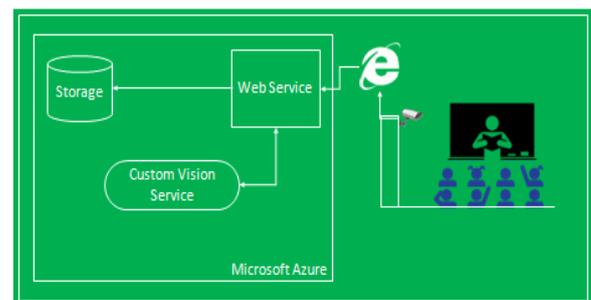
1. Introduction

The project aims towards the assistance of teachers at the time of taking attendance. The system solely focuses on face detection and recognition. The tools used to device the system are API's offered by Python 3.6, Open CV (for detection) and a few cognitive tools provided by Azure. The basic idea behind the project is face recognition linked to a database backend. The information of the student attending the class is stored here. The entire attendance is associated with two types of time stamps incorporated at the server end. The time stamp helps to keep a track of the hour conducted and the time for which number of people attended the class. Exceptions in the time stamp would be incorporated in order to cater for the students leaving the class or trying to bunk the class. In case of further exceptions the time stamp will be a scope for further development of the system. All queries or conditions of the students will be answered by the system on communication with the admin. If the admin finds that the system was at fault then it can always be fixed by the admin for smooth functioning of the class monitoring system. Humans are smart and evolution has been something that has stayed constant all throughout the years. Since then we have started off on another quest of ours, the creation of an artificial intelligence. Humans are trying to create an apprentice who can help us achieve better efficiency. This project mainly focuses towards creating assistance for the teachers to help them avoid the inconvenience of calculating student's attendance. The project required a higher processing speed so we used Azure to run the code using azure shell with the free Azure account.

2. Dependencies

The requirement and dependencies of the project are:

- Python version 3.5 (latest versions can be used)
 - Anaconda 3 (for creating virtual environment)
 - Python Open CV is cv2 API.
 - The API 'pip' is used for installing different python API.
 - The API used in the project: Numpy, Scipy, Pandas, Tensorflow which uses keras as the backend.
 - The Tensorflow Gpu 1.2.1 or higher version which uses the protobuf compiler for creating a CNN (Convolution Neural Networks)
 - The initial approach of the project was started in an Ubuntu 14.01 but had to shift to windows as we encountered some driver issues with the webcam.
 - CUDA9.0 with cuDNN7.0.(1.7.0) py36\GPU\cuda91 cudnn70ss eVS207 15.49.1.85.3/7.1.2 x86_64 Python3.6/Compute3.0)
- Block Diagram (for real time application)



3. Datasets

The data sets created for the recognition were taken from live feed with the help of a webcam. The dataset created were 30 clips of the face to create the training set for the existing model. Based on the training set the recognition of the faces takes place. The samples are actually the cross validation set. The pre trained data set was obtained from the Face Net in the davidsandbergs repository in the Github. We also referred thesharren7 repository in Github for the real time face detection. Inception-resnet-v1Casia –webface 7. A few examples are given below.

4. Face Detection

Comparison of Haarcascade and CNN

The initial stage of the project we tried to detect faces using Open CV. The python API for Open CV is imported as cv2 which can easily be installed using pipthe classifier required for the detection is known as the HaarCascade classifier which has pre-trained dataset stored in an XML file. It has different XML files for Eye Detection, Side Face Detection, Frontal Face Detection and detecting different objects using different XML files. Before the detection to start we need to pre process the data using BGR2Gray function which helps to convert the hue image to a gray image which is then detected based on the multi-scale classifier which fetches the pre-trained data from the XML files. The pre-trained data is then computed with the live feed that we receive from the computer web cam. The live feed which we fetch is then converted to frames for detection. This was the prior effort made to detect faces but on the later stage we used the CNN for detection. The differences that we observed were phenomenal. The draw backs wereinefficient response time of about 2sec.S0 the alternative method was using CNN.



Fig. 1: a) BGR2Gray b) Face detection

So we can see the differences in both Haarcascade as well as CNN in the next picture. The project uses an input layer, a single hidden layer for gray image conversion and a single layer for detecting the face, the project used face net by Google for the detection and later for recognition. In the next image we can observe the difference between the face box of Haarcascade and CNN. CNN creates a larger frame compared to a Haarcascade and a better efficiency when we have more number of frames to be detected. Using Haarcascade gives flickers and delays in the program response.CNN offer a higher processing power.

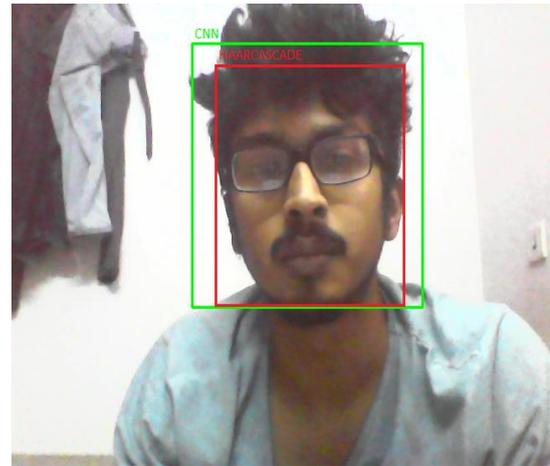


Fig. 2: CNN and Haarcascade Comparison

5. Neural Networks

The convolution neural network is a concept which is derived from the behaviour of neurons .The concept of training the neurons to make a specialised function. Similarly we come across different types of neural network architectures like:

- ANN(Artificial Neural Network)
- CNN(Convolution Neural Network)
- KNN(Kernel Next Neighbour)

Neural Network and Machine Learning

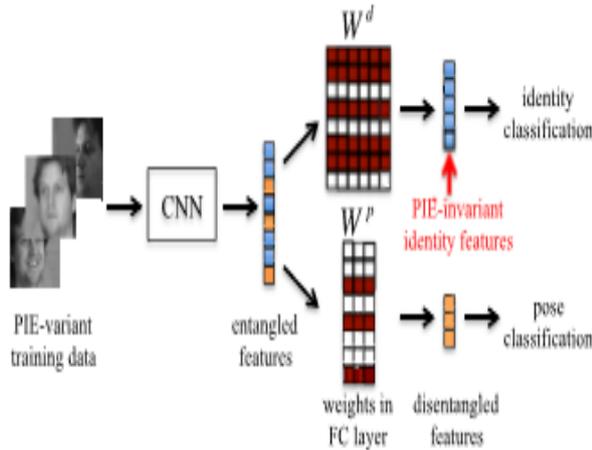
Similarly we make the neural networks using functions and we train the neurons using data set. Each data set is like a memory which makes the neurons stronger or more efficient with every iteration of the code. So the more number of data set means the more number of training experiences. The data sets are divided into three categories based upon the Machine Learning concept of Deep learning. The neurons can be trained for the purpose of applying Machine Learning concept for training the data set.They are training set, cross validation set and test set. So the concept of training a neuron is totally dependent upon the training set. The more number of training sets we have, the lesser chances of bias and over fitting. Bias occurs when the training set plotted against cross validation and test set andthe cross validation set begins to drift apart from the test set. This means the learning curve and cross validation curve

do not agree and the course of output is not desirable. Over fitting occurs when the test curve and the cross validation curve have no regressions between them, which means that there are no errors and the efficiency is 100%. But in such cases when more test sets are introduced then the learning curve takes a different course. Over fitting makes the neural network unsusceptible to any variation. The idea of neural network is to keep the functioning as generalised as possible to process ambiguous variations in the data set.These are the concepts of Machine Learning another branch of A.I. Based on the parameters the accuracy of a model can be calculated.

CNN Recognition

The project incorporates CNN for detection. The first is conversion of the image into a matrix .The matrix is then broken into smaller fragments of matrices and a weight is calculated. This is stored as a feature for the recognition purpose or the detection purpose depending upon the task assigned to the neural net. Based on the weights the recognition process work linked to a data base. The database is incorporated with the time stamp.

layer of the MTCNN the resolutions increase as follows: (3 X 3 X 1/96), (3 X 3 X 1/128), (3 X 3 X 1/256) and (3 X 3 X 1/320). The resolution increases but the size of the matrix decreases as with every progressing layer the detection of complex shapes and curves take place.



Till the time it reaches the fully connected Layer it is capable of identifying complicated shapes like the face. The features are created at the time of building classifiers. To analyse the model better we can make use of the confusion matrix created for understanding the different states of the algorithm. It defines steps like false positive, false negative, true positive and true negative. The confusion matrix for example can be given as:

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

Fig. 7: Confusion Matrix

Recognition and Time Stamp

The attendance system is directly associated with the recognition and the time stamp. When the face is recognised by the system a conditional flag is raised. This exception raised is a signal for the time stamp to get activated. The time stamp remains active till the time the face is present in the active frame. As soon as the person the ace is removed from the frame the time stamp is paused. The attendance of the person will be marked on the basis of duration for which the person is present in the class. If the duration falls below 60% of the duration of the class the person will be marked present for the session. This project will help the teachers in saving the time lost in taking the attendance of the students and hence they can utilise that time in taking the class.



Fig. 8: Time stamp with recognition

References

- [1] Face detection in colour images Rein-Lein Hsu, Student member IEEE, Mohamed Abdel Mottaleb, Member, IEEE and Anilk.Jain, Fellow, IEEE
- [2] Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning Hoo-Chang Shin, Member, IEEE, Holger R. Roth, Mingchen Gao, Le Lu, Senior Member, IEEE, Ziyue Xu, Isabella Noguez, Jianhua Yao, Daniel Mollura, and Ronald M. Summers*
- [3] Python.org & Opencv.org
- [4] Andrew Ng Machine Learning
- [5] Udacity Nanodegree Machine Learning Foundation
- [6] P Viola, M. Jones, "Robust, "Real time object detection", Compaque CRL,2001
- [7] Philip Ian Willson, John Fernandez "Facial Feature Detection using Haar Classifiers", Journal of Computing Sciences in colleges, Vol.21,no.4,pp.127-133, April 2006.
- [8] <https://developer.nvidia.com/cuda-zone>
- [9] <https://developer.nvidia.com/cuda/cudnn>
- [10] <https://www.tensorflow.org/>
- [11] Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition Xi Yin and XiaomingLiu, Member, IEEE