PRECOGNITION

An IP Camera Based Intelligent Attendance System Using Computer Vision

Md. Zillur Rahman

S.M.Foysal Ahmed

29.10.2019

North South University

ABSTRACT

Automatic face recognition (AFR) technologies have made many improvements in the changing world. Smart Attendance using Real-Time Face Recognition is a real-world solution which comes with day to day activities of handling student attendance system. Face recognition-based attendance system is a process of recognizing the students face for taking attendance by using face biometrics based on high - definition monitor video and other information technology. In my face recognition project, a computer system will be able to find and recognize human faces fast and precisely in images or videos that are being captured through a surveillance camera. Numerous algorithms and techniques have been developed for improving the performance of face recognition but the concept to be implemented here is Deep Learning. It helps in conversion of the frames of the video into images so that the face of the student can be easily recognized for their attendance so that the attendance database can be easily reflected automatically.

INTRODUCTION

Automated Attendance System (AAS) is a process to automatically estimate the presence or absence of the student in the classroom by using face recognition technology. It is also possible to recognize whether the student is sleeping or awake during the lecture and it can also be implemented in the exam sessions to ensure the presence of the student. The presence of the students can be determined by capturing their faces on to a high-definition monitor video streaming service.

The two common Human Face Recognition techniques are 1. Feature-based approach 2. Brightness-based approach. The Feature-based approach also known as local face recognition system, used in pointing the key features of the face like eyes, ears, nose, mouth, edges, etc., whereas the brightness-based approach also termed as the global face recognition system, used in recognizing all the parts of the image.

BACKGROUND

This project requires face detection, recognition and verification algorithms. There is a vast corpus of face verification and recognition [8] works. Reviewing it is out of the scope of this paper, so we will only briefly discuss the most relevant recent work. The works of [1, 2, 3] all employ a complex system of multiple stages, that combines the output of a deep convolutional network with PCA for dimensionality reduction and an SVM for classification.

Zhenyao et al. [3] employ a deep network to “warp” faces into a canonical frontal view and then learn CNN that classifies each face as belonging to a known identity. For face verification, PCA on the network output in conjunction with an ensemble of SVMs is used.

Taigman et al. [7] propose a multi-stage approach that aligns faces a general 3D shape model. A multi-class network is trained to perform the face recognition task on over four thousand identities. The authors also experimented with a so-called Siamese network where they directly optimize the L1-distance between two face features. Their best

performance on LFW (97.35%) stems from an ensemble of three networks using different alignments and colour channels. The predicted distances (non-linear SVM predictions based on the χ 2 kernel) of those networks are combined using a non-linear SVM.

Sun et al. [4, 1] propose a compact and therefore relatively cheap to compute network. They use an ensemble of 25 of these networks, each operating on a different face patch. For their final performance on LFW (99.47% [1]) the authors combine 50 responses (regular and flipped). Both PCA and a Joint Bayesian model [7] that effectively correspond to a linear transform in the embedding space are employed. Their method does not require explicit 2D/3D alignment. The networks are trained by using a combination of classification and verification loss. The verification loss is similar to the triplet loss we employ [5, 6], in that it minimizes the L2-distance between faces of the same identity and enforces a margin between the distance of faces of different identities. The main difference is that only pairs of images are compared, whereas the triplet loss encourages a relative distance constraint.

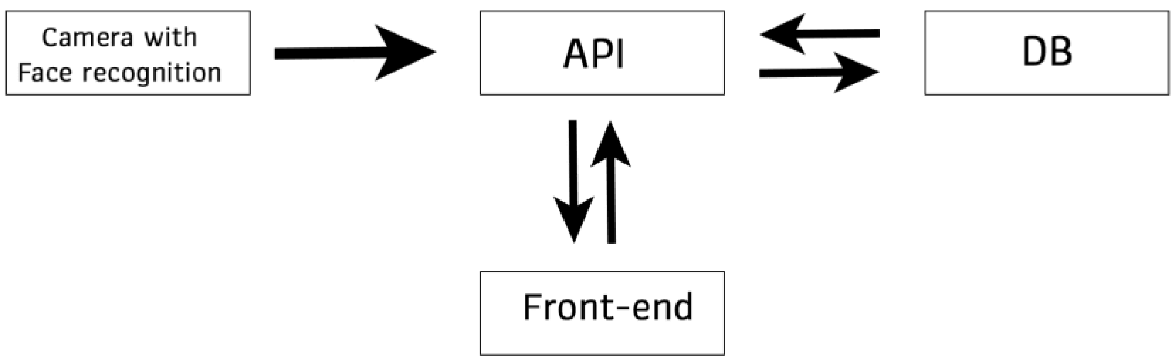
OBJECTIVES & SCOPES

Most offices, educational institutions and government buildings use biometric systems for logging in work hours and restricting access to only those people that are permitted. Although facial recognition is now being implemented in many countries for this purpose, they still largely rely on fingerprint sensors because of the lower chance of duplication and fraud. The downside to this is that recording a fingerprint is a task that requires relatively more effort.

By far the most interesting implementation of facial recognition in attendance systems is that of event attendance. The deployment of this system poses many benefits, especially for large events. Currently most events such as large corporate conferences and music concerts still largely rely on tickets, passes, and scanning of barcodes.

When audience numbers hit the late hundreds and even the thousands, the inefficiency of current systems begin to show. Checking tickets and scanning barcodes is a process that can take time depending on numerous variables.

PROCEDURE



1. **Capture video:**

   The Camera is fixed at a specific distance inside a classroom to capture videos of the frontal images of the entire students of the class.

2. **Separate as frames from the video:**

   The captured video needs to be converted into frames per second for easier detection and recognition of the students.

3. **Face Detection:**

   Face Detection is the process where the image, given as an input (picture) is searched to find any face, after finding the face image processing cleans up the facial image for easier recognition of the face. CNN algorithm can be implemented to detect the faces.

4. **Face Recognition:**

   After the completion of detecting and processing the face, it is compared to the faces present in the students' database to update the attendance of the students.

5. **Post-Processing:**

   The post-processing mechanism involves the process of updating the names of the student into an excel sheet. The excel sheet can be maintained on a weekly basis or monthly basis to record the students' attendance. This attendance record can be sent to parents or guardians of students to report the performance of the student.

BUDGET

SL	COST FACTORS	AMOUNT
1.	IP CAMERA	3000+
2.	HOSTING	?+
	TOTAL=	3000++

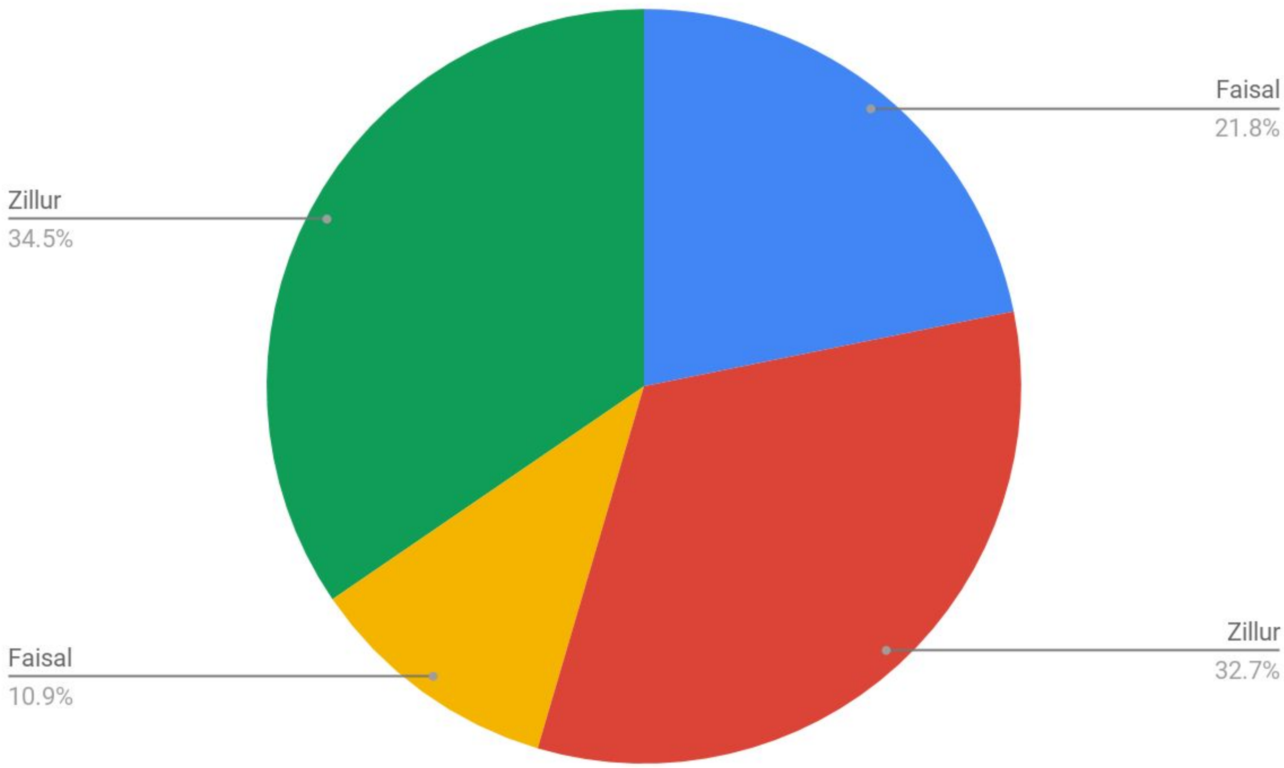
LANGUAGE, LIBRARIES & FRAMEWORKS

1. Python
2. FaceNet library
3. Numpy
4. Tensorflow
5. Keras
6. Flask
7. React
8. Mysql

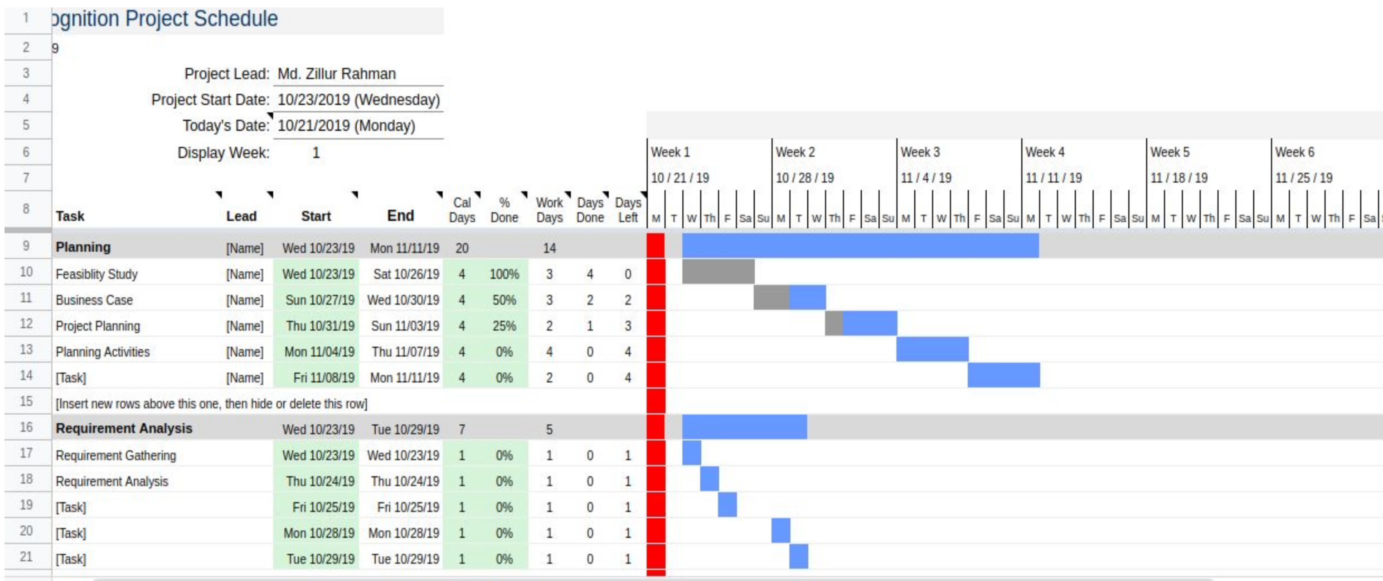
TASK LIST

SL	DESCRIPTION
1.	Environment Set Up
2.	Identify Face and its location from a given photo by using library
3.	Calculate Facial Parametres and save it to numpy array and save this so-called "encoded" data by linking them to a name so that they can be compared with a future picture
4.	Connect IP camera using open cv and feed the video as frames to the model
5.	Identify the face of unnamed photo and repeat task 2 and 3
6.	Automate task 2 and 4 for ‘n’ numbers of people
7.	Build Flask API to receive the data from our face recognition model and redistribute it to the front when requested.
8.	Build user interface (Front-end) using React
9.	Testing
10.	Deployment in Linux server

TASK DISTRIBUTION



GANTT CHART



ɔgnition Project Schedule

9

Project Lead: Md. Zillur Rahman

Project Start Date: 10/23/2019 (Wednesday)

Today's Date: 10/21/2019 (Monday)

Display Week: 1

Week 1 10 / 21 / 19 | Week 2 10 / 28 / 19 | Week 3 11 / 4 / 19 | Week 4 11 / 11 / 19 | Week 5 11 / 18 / 19 | Week 6 11 / 25 / 19

Task	Lead	Start	End	Cal Days	% Done	Work Days	Days Done	Days Left
Planning	[Name]	Wed 10/23/19	Mon 11/11/19	20		14		
Feasiblity Study	[Name]	Wed 10/23/19	Sat 10/26/19	4	100%	3	4	0
Business Case	[Name]	Sun 10/27/19	Wed 10/30/19	4	50%	3	2	2
Project Planning	[Name]	Thu 10/31/19	Sun 11/03/19	4	25%	2	1	3
Planning Activities	[Name]	Mon 11/04/19	Thu 11/07/19	4	0%	4	0	4
[Task]	[Name]	Fri 11/08/19	Mon 11/11/19	4	0%	2	0	4
[Insert new rows above this one, then hide or delete this row]								
Requirement Analysis		Wed 10/23/19	Tue 10/29/19	7		5		
Requirement Gathering		Wed 10/23/19	Wed 10/23/19	1	0%	1	0	1
Requirement Analysis		Thu 10/24/19	Thu 10/24/19	1	0%	1	0	1
[Task]		Fri 10/25/19	Fri 10/25/19	1	0%	1	0	1
[Task]		Mon 10/28/19	Mon 10/28/19	1	0%	1	0	1
[Task]		Tue 10/29/19	Tue 10/29/19	1	0%	1	0	1

EXPECTED OUTCOMES

This project is implemented using the the very fundamental functionalities of computer vision library. But this project will be a boost for the people who want to explore the ocean of computer vision and its future applications both in academic and research fields. Moreover this whole project can be built as a library which can be integrated to any software system opening the door of portability and ease.

REFERENCES

1. Y. Sun, X. Wang, and X. Tang. Deeply learned face representations are sparse, selective, and robust. CoRR, abs/1412.1265, 2014. 1, 2, 5, 8
2. Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deep face: Closing the gap to human-level performance in face verification. In IEEE Conf. on CVPR, 2014. 1, 2, 5, 8
3. Z. Zhu, P. Luo, X. Wang, and X. Tang. Recover canonical view faces in the wild with deep neural networks. CoRR, abs/1404.3543, 2014. 2

4. Y. Sun, X. Wang, and X. Tang. Deep learning face representation by joint identification-verification. CoRR, abs/1406.4773, 2014. 1, 2, 3
5. M. Schultz and T. Joachims. Learning a distance metric from relative comparisons. In S. Thrun, L. Saul, and B. Schölkopf, editors, NIPS, pages 41–48. MIT Press, 2004. 2
6. K. Q. Weinberger, J. Blitzer, and L. K. Saul. Distance metric learning for large margin nearest neighbour classification. InNIPS. MIT Press, 2006. 2, 3
7. D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun. Bayesian face revisited: A joint formulation. In Proc. ECCV, 2012. 2
8. Florian Schroff, Dmitry Kalenichenko, James Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815-823