A Project Mid-term defence for the Degree of Bachelor of Science in Computer and Information Technology

FACE RECOGNITION ATTENDANCE SYSTEM USING MTCNN

Submitted for the partial fulfilment of the requirement for the degree of Bachelor of Science in Computer and Information Technology

Sabina Shrestha
Abhisesh Chapagain
Dipika Timalsina

Swastik College

Department of Computer Science and Information Technology

Institute of Science and Technology

Tribhuvan University

February, 2023

DECLARATION

This project, entitled" FACE RECOGNITION ATTENDANCE SYSTEM USING MTCNN" is based on our own research. Other researchers' related work on this issue has been recognized. We are responsible for all liabilities pertaining to the data correctness and validity, as well as any other material contained herein

.....
Abhisesh Chapagain
.....
Sabina Shrestha
.....
Dipika Timalsina

Recommendation

I hereby recommend that the project prepared under my supervision by Mr. Sagar Rana Magar entitled " FACE RECOGNITION ATTENDANCE SYSTEM USING MTCNN "be acceptedas fulfilling in partial requirement for the degree of Bachelor of Science in Computer Science and Information Technology. In my best knowledge, this is an original work in Computer Science and Information Technology.

.....
Mr. Sagar Rana Magar
Supervisor
Lecturer
Swastik College
Chardobato, Bhaktapur

CERTIFICATE

This is to certify that this project prepared by ABHISESH CHAPAGAIN, SABINA SHRESTHA and DIPIKA TIMALSINA entitled " FACE RECOGNITION ATTENDANCE SYSTEM USING MTCNN" in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Science and Information Technology has been well studied. In our opinion, it is satisfactory in scope and quality as a project for the required degree.

.....
Mr. Sagar Rana Magar
Supervisor
Lecturer
Swastik College Chardobato,
Bhaktapur

.....
External Examiner

.....
Mrs. Sristi Khatiwada
Co-ordinator
Swastik College
Chardobato, Bhaktapur

Acknowledgement

A report entitled “FACE RECOGNITION ATTENDANCE SYSTEM USING MTCNN" has been prepared to fulfill the requirement of the undergraduate project. The report is deeply indebted to Tribhuvan University and Swastik College for incorporating a project report in the bachelor program which provides a wonderful opportunity to carry out practical knowledge on any topic.

Our forthright gratefulness goes to Mrs. Sristi Khatiwada, our program coordinator, and Mr. Sagar Rana Magar, Supervisor, Department of Computer Science and Information technology (B.Sc.CSIT) for your coordinating and supporting this project.
Lastly, we are very thankful to all the people who helped us with this project. We humbly and gratefully acknowledge the support of our parents and teachers for their immense support. We also want to thank our friends for their input in the project.

Thanking you,
Sabina Shrestha
Dipika Timalsina
Abisesh Chapagain

Abstract

In this era of Artificial Intelligence, manually calling a student and taking attendance is a time consuming task. Hence, we have proposed an automated student attendance system based on face recognition. Face recognition systems are very useful in life applications especially in security control systems. In our proposed approach, firstly, video framing is performed by activating the camera through a userfriendly interface. The face is detected and segmented from the video frame by using MTCNN algorithm. In the face recognition stage, we have used facenet pre-trained and used transfer learning for new faces. Once the face embeddings are extracted from the model, SVM classifier is used to classify and recognize faces into different classes. Finally, the attendance of the recognized student will be marked and saved in the dataset and can be viewed according to the date. All these features can be accessed through flutter applications. In addition, new student images can be added and trained within the app. This project intends to serve as an efficient substitute for traditional manual attendance systems.

Keywords: Face Detection, Algorithms, Features extraction, FaceNet , SVC, MTCNN

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AI(Artificial Intelligence)
ML (Machine Learning)
SVM (Support Vector Machine)

Chapter1:Introduction

1.1Background

The main objective of this project is to develop a face recognition-based automated student attendance system. The system detect and recognize face of students in the frame of camera, take attendance and store the result in the database for future reference. The students have to register in the dataset to be recognized. The student enrolment can also be done on the spot through a user-friendly interface.

We recognize the face of family, friends or someone we are familiar with. Human mind takes several steps to finally recognize the face. Human intelligence works in a way, it first allows us to collect the information and interpret the information in recognition process. According to Robinson-Riegler, G., Robinson-Riegler, B. (2008) after the processing of the image is done by the human visual system, we then classify shape, size, contour and texture of the object to analyze the information. The analyzed information is then compared to other representations of faces that is in the memory to recognize. Since, it is not that easy to build an automated system having the visual capacity as high as human. However, we need a huge amount of memory to store abundant amount of face characters and examine on them. Thus, in order to overcome the human limitations, computer system with almost infinite memory, very high processing speed and power are used for facial recognition.

The human face is a yet another artistic expression identity. Face recognition is thus described as a bio-metric approach for identifying an individual by comparing real-time capture images with images stored in that person's database (Margaret Rouse, 2012).

Face recognition came into existence from 1960. Woody Bledsoe, Helen Chan Wolf and Charles Bisson had introduced a system that required the administrator to locate eyes, ears, nose and mouth for images.The common reference point and the distance and ratio between located features are calculated and compared. The studies on Facial Recognition were further carried out by Goldstein, Harmon, and Lesk in 1970 adding further features like hair color lip thickness to automate the recognition and increase the accuracy. Kirby and Sirovich proposed principle component analysis (PCA) to address the face recognition problem for the first time in 1988. Many facial recognition research have been undertaken since then, and they are still ongoing today (Ashley DuVal, 2012).

1.2 Problem Statement

Attendances of every student are being maintained by every school, college and university. Evidences have shown that there is a significant correlation between students' attendances and their academic performances. There was also a claim stated that the students who have poor attendance records will generally link to poor retention. Therefore, faculty has to maintain proper record for the attendance. The manual attendance record system is not efficient and requires more time to arrange record and to calculate the average attendance of each student. Hence there is a requirement of a system that will solve the

problem of student record arrangement and student average attendance calculation. One alternative to make student attendance system automatic is provided by facial recognition.

Time and again, traditional attendance marking system faces a numerous problems, and it is time consuming as well. Face Recognition Student attendance system simply minimize the problems faced in Classical system like calling the name of the student or checking respective identity cards. These inefficient method not only disturbs the teaching process but also distracts the student during the examinations time. Furthermore, attendance sheet is passed around the classroom during the lecture session. If the classroom is very big and if the number of student is also large, then there might be a case where the attendance of absents student is done by his colleague and also it might be difficult to pass the attendance sheet all around the class.As a result, a face recognition student attendance system is proposed to replace the manual signing of students' presence, which is inconvenient and causes students to have become distracted while signing for their attendance. Furthermore, using Face Recognition System fraudulent approach can be removed and the burdensome of calling the name of every students several times to ensure the presence of student can be minimized.

1.3 Project Objectives

The prime objective of this project is to build a face recognition based Student Attendance System and to record the attendance of identified students.

1.4 Scope and Limitation

Scope and Limitation of Text Summarization are listed below:

Scope

Face Recognition is becoming more prominent in our society. It has major progress in the field of security and keeping the track of the people working in any organization or the presence/absence record of students in Educational Institution. Since, the traditional Attendance System in our College is manual and it is time consuming and ineffective as well. Thus, this project aims to automate the traditional system where the attendance is marked manually. Here, through our App, any teacher can easily take an attendance of whole class on a single tap. Furthermore, this system minimizes the fraudulent marking of any student through the actual face recognition. Using this system any corporate offices, school and organization can replace their traditional way of maintaining attendance of the employees and can also generate their availability(presence) report throughout the month.

Limitation

Small face region gives inaccurate features, this will decrease the performance of the system. The laptop built in webcam is the default device in this proposed approach to capture images. The webcam and lighting source of the laptop have low performance which cause the captured images appear to be darker and blurred. This causes the system to function only the best if the test image and train image are both captured at the same place under approximately the same illumination. Besides, false recognition occurs when the facial image is blurred. The blurred image caused by the after image created by movement will

degrade the performance. The face feature extracted from the blurred image would be totally different compared to the train image resulting in false recognition. In addition, if an individual wears makeup in the image for face recognition, the important features will be covered. Similarly, face region should not be covered by hair, beard or any accessories to ensure better performance. For instance, a girl provides a facial image with her face covered by hair, it causes false recognition to occur if the girl ties her hair. This is because anything covering the face region will be assumed as a face feature. This causes a relatively large difference between test image and train image. Different levels of brightness or lighting could be a challenging problem for face recognition.

1.5 Development Methodology

For this project, we use the Rapid Application Development (RAD) model as a development methodology. RAD is the most effective form of app development that anchors itself on quick delivery, reusable blocks of code, feedback-based development, and rapid prototyping, and Incremental progress is visible and measurable, that's the way we use this model in our project.

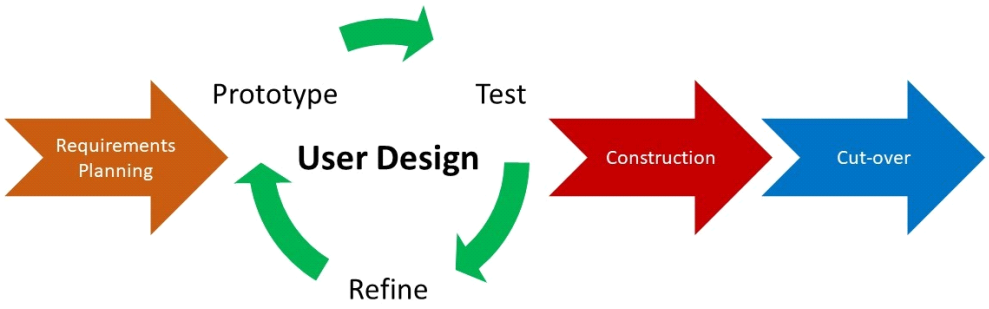

Figure 1.1: Rapid Application Development Model

1.6 Report Organization

The organization of a report on face recognition using MTCNN typically includes the following sections:
The report organization of each chapter that have on the report it arranges will refer to the specific format and it easy for readers to understand the whole of report. The report is started with chapter 1 that explains about introduction, problem statement, objectives, scope and

limitations and methodology. The next chapter 2 explains about the literature review related the paper research for the system development. Then, chapter 3 discussed about the system analysis where requirement analysis and feasibility analysis are required in system development. Then, chapter 4 deals the project design and modelling are the core part of the development process. Then, chapter 5 discussed about the implementation and testing of the system. Lastly, chapter 6 conclude on the system development. Chapter7 this section lists all the sources cited in the report.

Chapter 2: Literature Review

2.1 Background Study

2.1.1 Student Attendance System

Arun Katara et al.(2017) mentioned disadvantages of RFID (Radio Frequency Identification) card system, fingerprint system and iris recognition system. RFID card system is implemented only because of its simplicity. However, the problem with this system is that the students tends to help their friend to mark their attendance as long as they have their friend's id card. The fingerprint system is indeed effective but not efficient because it takes time for the verification process so the user has to line up and perform the verification one by one. However for face recognition, the human face is always exposed and contains less information compared to iris. Voice recognition is available, but it is less accurate compared to other methods. Hence, a face recognition system is suggested to be implemented in the student attendance system.

2.1.2 Face Detection

Differences between face detection and face recognition are often misunderstood. Face detection is to determine only the face segment or face region from the image, whereas face recognition is to identify the owner of the facial image. Sanjana devi et al. (2017) and WeiLun Chao (2007) presented a few factors which cause face detection and face recognition to encounter difficulties. These factors consist of background, illumination, pose, expression, occlusion, rotation, scaling and translation.

- Background - Variation of background and environment in the video frame affects the face recognition efficiency.
- Illumination - Variation in the lighting of the environment also alters the facial detection and recognition.
- Pose - variation means the different angle of the acquired facial images. This change in the pose causes distortion to the recognition process, especially for Eigen face recognition.

Some of the Face Detection Algorithms so far are:

- Viola Jones Algorithm
- AdaBoost Algorithm
- Neural-Network

Neural network is initially used only in face detection. It is then further studied to be imple- mented in face recognition. Network (ANN) was studied for face recognition. Convolutional Neural (CNN) is another neural network algorithm for face detection. The most widely used algorithm in face detection is MTCNN (abbreviation for Multi-task Cascaded Convolutional Networks). The MTCNN algorithm is a face detection and face alignment method based on deep learning. It can complete the tasks of face detection and face alignment at the same time. Compared with traditional algorithms, it has better performance and faster detection speed.

2.1.3 Feature Extraction

A feature is a set of data that represents the content of digital images. Face recognition re- quires the extraction of facial features. The selection of features, on the other hand, could be a difficult task. To achieve high accuracy, the feature extraction algorithm must be reliable and robust throughout a wide range of changes. There are a few feature extraction methods for face recognition. Some of them are as follows:

- PCA (Principal Component Analysis)

  In the paper of Bhuvaneshwari et al. (2017), Abhishek Singh and Saurabh Kumar (2012) and Liton Chandra Paul and Abdulla Al Sumam (2012), they proposed PCA for the face recognition. D. Nithya (2015) also used PCA in face recognitionbased student attendance system. PCA is famous with its robust and highspeed computation. Basically, PCA retains data variation and remove unnecessary existing correlations among the original features. PCA is basically a dimension reduction algorithm. It compresses each facial image which is represented by the matrix into single column vector. Furthermore, PCA removes average value from image to centralize the image data. The Principle Component of distribution of facial images is known as Eigen faces. Every single facial image from training set contributes to Eigen faces. As a result, Eigen face encodes best variation among known facial images.

- LDA (Linear discriminant analysis)

LDA also known as Fisher face is another popular algorithm for face recognition. In the paper by Suman Kumar Bhattacharyya and Kumar Rahul (2013), LDA was proposed for face recognition. LDA extract features by grouping images of the same class and separate images of

different classes. LDA is able to perform well even with different facial expressions, illumination and pose d0ue to its class separation characteristic. Same class is defined by facial images of the same individual, but with different facial expressions, varying lighting or pose, whereas facial images of person with different identity are categorized as different classes. Same class images yield within-class scatter matrix meanwhile different class images yield between-class scatter matrix. LDA manage to maximize the ratio of the determinant of the between-class scatter matrix over the determinant of the within class scatter matrix.

- LBP (Local Binary Patterns)

The original LBP (Local Binary Patterns) operator was introduced by the paper of Timo Ojala et al. (2002). In the paper by Md. Abdur Rahim et al. (2013), they proposed LBP to extract both texture details and contour to represent facial images. LBP divides each facial image into smaller regions and histogram of each region is extracted. The histograms of every region are concatenated into a single feature vector. This feature vector is the representation of the facial image and Chi square statistic is used to measure similarities between facial images. The smallest window size of each region is 3 by 3. It is computed by thresholding each pixel in a window where middle pixel is the threshold value. The neighborhood larger than threshold value is assigned to 1 whereas the neighborhood lower than threshold value is assigned to 0. Then the resulting binary pixels will form a byte value representing center pixel.

- FaceNet

FaceNet is the name of the facial recognition system that was proposed by Google Researchers in 2015 in the paper titled FaceNet: A Unified Embedding for Face Recognition and Clustering. It achieved state-of-the-art results in the many benchmark face recognition dataset such as Labeled Faces in the Wild (LFW) and Youtube Face Database. They proposed an approach in which it generates a high-quality face mapping from the images using deep learning architectures such as ZF-Net and Inception. Then it used a method called triplet loss as a loss function to train this architecture.

2.1.4 Feature Classification and Face Recognition

Classification involves the process of identification of faces. Deepesh Raj (2011) mentioned several types of distance classifiers such as Euclidean Distance, City Block Distance and Mahalanobis distance for face recognition. Vladimir N. Vapnik and Alexey Ya. Chervonenkis(1963) suggested a supervised learning algorithm that solves complex optimization problem that maximizes the margin and where the constraints say that points of each category should be on the correct side of the hyperplane. The algorithm is known as Support Vector Machine (SVM). The goal is: Finding a hyperplane with the maximum margin (margin is basically a protected space around hyperplane equation) and algorithm tries to have maximum margin with the closest points (known as support vectors).

In other words, “The goal is to maximize the minimum distance.” for the distance: euclidean distance between data point and the hyperplane.

2.2 Literature Review

Authors in [3] proposed a model of an automated attendance system. The model focuses on how face recognition incorporated with Radio Frequency Identification (RFID) detect the authorized students and counts as they get in and get out form the classroom. The system keeps the

authentic record of every registered student. The system also keeps the data of every student registered for a particular course in the attendance log and provides necessary information according to the need.

In [4], authors proposed an attendance system based on facial recognition. The algorithms like Viola-Jones and Histogram of Oriented Gradients (HOG) features along with Support Vector Machine (SVM) classifier were used to implement the system. Various real time scenarios such as scaling, illumination, occlusions and pose was considered by the authors. Quantitative analysis was done on the basis of Peak Signal to Noise Ratio (PSNR) values and was implemented in MATLAB GUI.

Authors in [5] researches to get best facial recognition algorithm (Eigenface and Fisherface) provided by the Open CV 2.4.8 by comparing the Receiver Operating Characteristics (ROC) curve and then implemented it in the attendance system. Based on the experiments carried out in this paper, the ROC curve proved that, Eigenface achieves better result than Fisherface. System implemented using Eigenface algorithm achieved an accuracy rate of 70% to 90%.

In [6], authors proposed a method for student attendance system in classroom using face recognition technique by combining Discrete Wavelet Transforms (DWT) and Discrete Cosine Transform (DCT). These algorithms were used to extract the features of students face followed by applying Radial Basis Function (RBF) for classifying the facial objects. This system achieved an accuracy rate of 82%.

In this paper [7], authors have designed and implemented an attendance system which uses iris biometrics. Initially, the attendees were asked to register their details along with their unique iris template. At the time of attendance, the system automatically took class attendance by capturing the eye image of each attendee, recognizing their iris, and searching for a match in the created database. The prototype was web based.

Chapter 3: System Analysis

3.1 System Analysis

System analysis for "Face Recognition Attendance System Using MTCNN " involves the examination of the problem of face recognition, identification of requirements, evaluation of feasibility, analysis of user needs.

3.1.1 Requirement Analysis

- Functional Requirements

Functional requirements refer to the specific functions or tasks that a system must be able to perform in order to meet the needs of its users. In the context of Face Recognition Attendance System Using MTCNN, the following are some of the key functional requirements:

- **Input Image**: We used our own database to design a real-time face recognition student attendance system in different conditions. The conditions included center-light, with glasses, happy, left-light The images captured by using a laptop built-in, without glasses, normal, right-light, sad, sleepy, surprised and wink. These different variations provided by the database are able to ensure the system to be operated consistently in a variety of situations and conditions. For our database, the images of students are captured by using a laptop built-in camera, mobile phone camera, and digital camera. The images captured by using a laptop built-in camera are categorized as low-quality images, whereas mobile phone camera and digital camera captured images are categorized as high-quality images. The image quality of each student ranges from low to high. The enrollment can be done on the spot through a user-friendly interface.
- **Preprocessing**: Testing set and training set images are captured using a camera. There is unwanted noise and uneven lighting exists in the images. Therefore, several preprocessing steps are necessary be- fore proceeding to feature extraction. Pre-processing steps that would be carried out include 160*160 pixels scaling of the detected face on the image, conversion of detected facial pixel into RGB channel and standarizing the pixel values in the range [0,1] using mean and standard deviation.
- **Model Training**: The system must be able to train the MTCNN-based model using large dataset of image data.
- **Result**: The system must be able to provide the accurate result as much as possible.

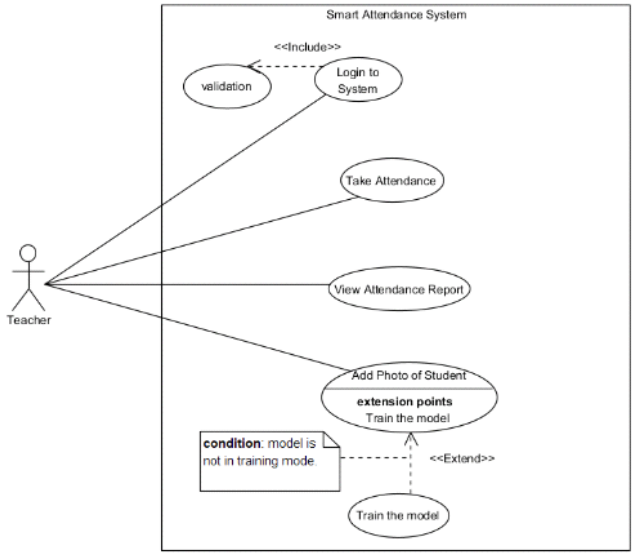

Figure 3.1: Use Case Diagram

- Non-Functional Requirements

Non-functional requirements are the attributes or qualities of a system that are not directly related to its functional capabilities, but are still essential for its overall effectiveness. In the context of Face recognition attendance system using MTCNN, the following are some of the key non-functional requirements:

- Accuracy: The system must be able to recognize face accurately.
- Speed: The system must be able to do attendance in real-time and be fast enough to meet the needs of its users.

3.1.2 Feasibility Analysis

Feasibility analysis is the process of evaluating the viability of a proposed solution. The objective of feasibility analysis is to determine whether the proposed solution is feasible and practical, and whether the benefits of the solution outweigh the costs and risks involved in its implementation. In the context of Face recognition attendance system using MTCNN, feasibility analysis involves evaluating the technical, operational, economic, and schedule feasibility of the proposed solution.

- Technical Feasibility: This text summarization web-application fully works on all system which supports internet connectivity.
- Schedule Feasibility: For each task of the project, proper estimation and splitting of the time have been done. Overall, the calculated time is sufficient enough to complete project in time.

Chapter4: System Design

4.1 Design: Structured Approach

4.2AlgorithmDetails

4.2.1MTCNN

Multi-task Cascaded Convolutional Networks (MTCNN) is a framework developed as a solution for both face detection and face alignment. It was proposed on the paper “Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks” by Zhang, Zhang and Zhifeng. The process consists of three stages of convolutional networks that are able to recognize faces and landmark location such as eyes, nose, and mouth.

Stage 1: The Proposal Network (P-Net) This first stage is a fully convolutional network (FCN). The difference between a CNN and a FCN is that a fully convolutional network does not use a dense layer as part of the architecture. This Proposal Network is used to obtain candidate windows and their bounding box regression vectors. Bounding box regression is a popular technique to predict the localization of boxes when the goal is detecting an object of some pre-defined class, in this case faces. After obtaining the bounding box vectors, some refinement is done to combine overlapping regions. The final output of this stage is all candidate windows after refinement to downsize the volume of candidates.

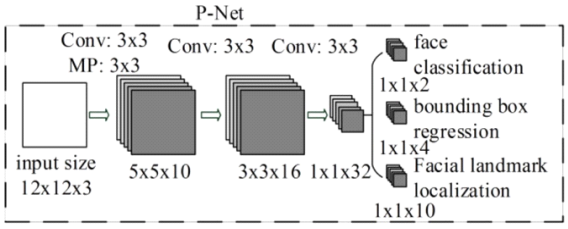

Figure 4.1: P-Net model for face detection

Stage 2: The Refine Network (R-Net) All candidates from the P-Net are fed into the Refine Network. Notice that this network is a CNN, not a FCN like the one before since there is a dense layer at the last stage of the network architecture. The R-Net further reduces the number of

candidates, performs calibration with bounding box regression and employs non-maximum suppression (NMS) to merge overlapping candidates.

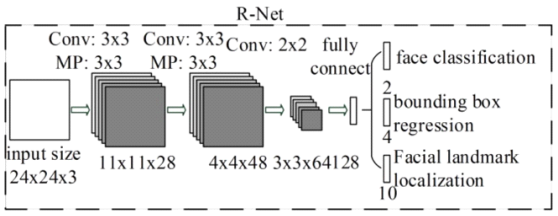

Figure 4.2: R-Net model for face detection

Stage 3: The Output Network (O-Net) This stage is similar to the R-Net, but this Output Network aims to describe the face in more detail and output the five facial landmarks' positions for eyes, nose and mouth.

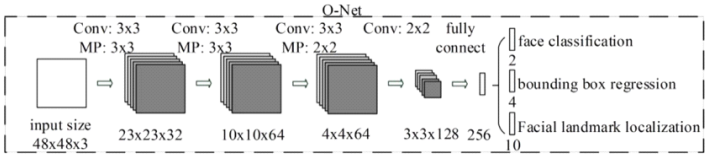

Figure 4.3: Q-Net model for face detection

4.2.2 FaceNet

FaceNet is the name of the facial recognition system that was proposed by Google Researchers in 2015 in the paper titled FaceNet: A Unified Embedding for Face Recognition and Clustering. They proposed an approach in which it generates a high-quality face mapping from the images using deep learning architectures such as ZF-Net and Inception. Then it used a method called triplet loss as a loss function to train this architecture.

Architecture

Figure 4.4: FaceNet Architecture

FaceNet employs end to end learning in its architecture. It uses ZF-Net or Inception as its under- lying architecture. It also adds several 1*1 convolutions to decrease the number of parameters. These deep learning models outputs an embedding of the image f(x) with L2 normalization performed on it. These embeddings then passed into the loss function to calculate the loss.

Triplet Loss

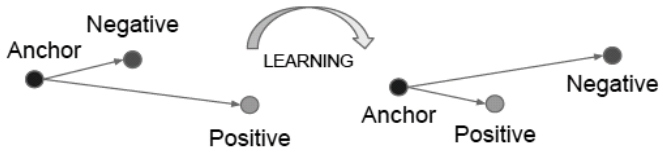

Figure 4.5: Triplet loss

The intuition behind triplet loss function is that we want our anchor image (image of a specific person A) to be closer to positive images (all the images of person A) as compared to negative images (all the other images). Triplet loss function can be formally defined as —

$$\sum_{i}^{N} [|f(x_i^a) - f(x_i^p)|_2 - |f(x_i^a) - f(x_i^n)|_2 + \alpha]$$

- x_i represents an image
- $f(x_i)$ represents the embedding of an image
- α represents the margin between positive and negative pair
- subscript (a, p and n) represents anchor, positive and negative images respectively.

Choosing the correct image pairs is extremely important as there will be a lot of image pairs that will satisfy this condition and hence our model won't learn much from them and will also converge slowly because of that. we need to select triplets such that $|f(x^a_i) - f(x^p_i)|^2$ is maxi- mum and $|f(x^a_i) - f(x^n_i)|^2$ is minimum.

Training This model is trained using Stochastic Gradient Descent (SGD) with backpropagation and AdaGrad. This model is trained on a CPU cluster for 1k-2k hours. The steady decrease in loss (and increase in accuracy) was observed after 500 hours of training. This model is trained using two networks:

• ZF-Net:

layer	size-in	size-out	kernel	param	FLPS
conv1	220×220×3	110×110×64	7×7×3, 2	9K	115M
pool1	110×110×64	55×55×64	3×3×64, 2	0	
rnorm1	55×55×64	55×55×64		0	
conv2a	55×55×64	55×55×64	1×1×64, 1	4K	13M
conv2	55×55×64	55×55×192	3×3×64, 1	111K	335M
rnorm2	55×55×192	55×55×192		0	
pool2	55×55×192	28×28×192	3×3×192, 2	0	
conv3a	28×28×192	28×28×192	1×1×192, 1	37K	29M
conv3	28×28×192	28×28×384	3×3×192, 1	664K	521M
pool3	28×28×384	14×14×384	3×3×384, 2	0	
conv4a	14×14×384	14×14×384	1×1×384, 1	148K	29M
conv4	14×14×384	14×14×256	3×3×384, 1	885K	173M
conv5a	14×14×256	14×14×256	1×1×256, 1	66K	13M
conv5	14×14×256	14×14×256	3×3×256, 1	590K	116M
conv6a	14×14×256	14×14×256	1×1×256, 1	66K	13M
conv6	14×14×256	14×14×256	3×3×256, 1	590K	116M
pool4	14×14×256	7×7×256	3×3×256, 2	0	
concat	7×7×256	7×7×256		0	
fc1	7×7×256	1×32×128	maxout p=2	103M	103M
fc2	1×32×128	1×32×128	maxout p=2	34M	34M
fc7128	1×32×128	1×1×128		524K	0.5M
L2	1×1×128	1×1×128		0	
total				140M	1.6B

Figure 4.6: ZF-Net Summary

• Inception:

type	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj (p)	params	FLOPS
conv1 (7×7×3, 2)	112×112×64	1							9K	119M
max pool + norm	56×56×64	0						m 3×3, 2		
inception (2)	56×56×192	2		64	192				115K	360M
norm + max pool	28×28×192	0						m 3×3, 2		
inception (3a)	28×28×256	2	64	96	128	16	32	m, 32p	164K	128M
inception (3b)	28×28×320	2	64	96	128	32	64	L_2, 64p	228K	179M
inception (3c)	14×14×640	2	0	128	256,2	32	64,2	m 3×3,2	398K	108M
inception (4a)	14×14×640	2	256	96	192	32	64	L_2, 128p	545K	107M
inception (4b)	14×14×640	2	224	112	224	32	64	L_2, 128p	595K	117M
inception (4c)	14×14×640	2	192	128	256	32	64	L_2, 128p	654K	128M
inception (4d)	14×14×640	2	160	144	288	32	64	L_2, 128p	722K	142M
inception (4e)	7×7×1024	2	0	160	256,2	64	128,2	m 3×3,2	717K	56M
inception (5a)	7×7×1024	2	384	192	384	48	128	L_2, 128p	1.6M	78M
inception (5b)	7×7×1024	2	384	192	384	48	128	m, 128p	1.6M	78M
avg pool	1×1×1024	0								
fully conn	1×1×128	1							131K	0.1M
L2 normalization	1×1×128	0								
total									7.5M	1.6B

Figure 4.7: Inception model summary

4.2.3 Linear SVC

The objective of SVC algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane.

Selecting best hyperplane

One reasonable choice as the best hyperplane is the one that represents the largest separation or margin between the two classes.

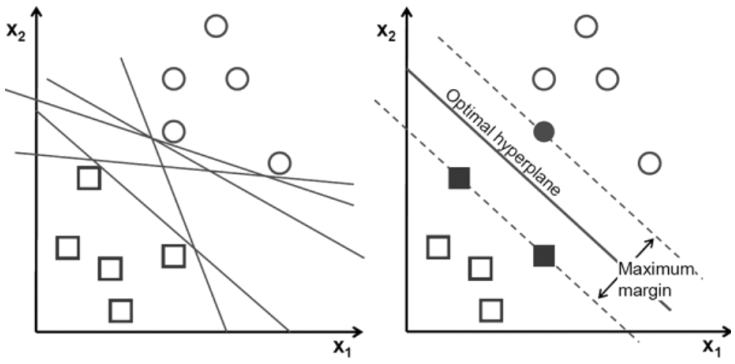

Figure 4.8: SVC: choosing best hyperplane

Chapter 5: Implementation and Testing

Implementing and Testing is executing a system to identify and fixed any gaps, errors, or missing requirements in contrary to the actual requirements. Software system testing is a process of verifying and validating a software application or program.

5.1 Implementation

5.1.1 Tools Used

- Flutter

Flutter is an open-source mobile application development framework created by google. It is a cross platform framework esp. used to develop applications for android and ios but can be extended to web platform and desktop. Components of Flutter includes - Dart platform - Flutter engine -Foundation library - Design-specific widgets

- Flask

For server side, Python framework Flask is used. Flask is micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. Flask supports extensions that can add application features as if they were implemented in Flask itself.

- TensorFlow

For implementation of summarization model, library Tensorflow was used. Tensorflow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is used in machine learning applications such as neural networks. Tensorflow was developed by the Google Brain team.

- MongoDB

For Database, MongoDB is used. MongoDB is a NoSQL database which stores the data in form of key-value pairs. It is an Open Source, Document Database which provides high performance and scalability along with data modelling and data management of huge sets of data in an enterprise application.

5.2 Result Analysis

5.2.1 Model Training and Validation Accuracy

We were able to collect face dataset from 20 of our friends which were divided into training and testing set with almost the ratio of 3:1. Our model showed the accuracy of 98.765 on train data and 97.114 on test data on facial data of students. Also f1 score for testing set were evaluated as shown in figure.

	precision	recall	f1-score	support
075BCT001	1.00	0.80	0.89	5
075BCT014	1.00	1.00	1.00	5
075BCT018	1.00	1.00	1.00	5
075BCT020	1.00	1.00	1.00	5
075BCT040	1.00	1.00	1.00	5
075BCT043	1.00	1.00	1.00	4
075BCT035	1.00	0.67	0.80	3
075BCT006	1.00	1.00	1.00	5
075BCT007	0.67	1.00	0.80	4
075BCT004	1.00	0.86	0.92	7
075BCT016	1.00	1.00	1.00	5
075BCT008	1.00	1.00	1.00	4
075BCT024	1.00	1.00	1.00	2
075BCT031	1.00	1.00	1.00	5
075BCT029	1.00	1.00	1.00	4
075BCT025	0.86	1.00	0.92	6
075BCT039	1.00	1.00	1.00	5
075BCT012	1.00	1.00	1.00	4
075BCT042	1.00	1.00	1.00	8
075BCT044	1.00	1.00	1.00	5
accuracy			0.97	96
macro avg	0.98	0.97	0.97	96
weighted avg	0.98	0.97	0.97	96

Figure 5.1: Accuracy, Recall, f1-score report

Chapter 6: Conclusion and Future Recommendations

6.2 Future Recommendations

Some more recommendation and enhancement on this project we could think of are:

- Make a student portal for uploading their images for dataset and viewing the attendance.
- The images of unknown people can be saved in an efficient manner and displayed in the system for better security.
- The number of training images can be reduced so that less storage is required. This can be done by removing duplicate images of the same person, or images with similar embeddings.
- A feature can be added where an student is automatically sent a warning if his attendance or working hours are below the threshold.

References

[1] Brownlee, J. (2019). Deep Learning for Computer Vision: Image Classi- fication, Object Detection, and Face Recognition in Python. Opgehaal van https://books.google.com.np/books?id=DOamDwAAQBAJ

[2]ttps://www.researchgate.net/publication/326261079_Face_detection_system_for_attendance_of_class_students

[3] Akbar, Md Sajid, et al. "Face Recognition and RFID Verified Attendance System." 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE). IEEE, 2018.

[4] Rathod, Hemantkumar, et al. "Automated attendance system using machine learning approach." 2017 International Conference on Nascen Technologies in Engineering (ICNTE). IEEE, 2017.

[5] Siswanto, Adrian Rhesa Septian, Anto Satriyo Nugroho, and Maulahikmah Galinium. "Implementation of face recognition algorithm for biometrics based time attendance system." 2014 International Conference on ICT For Smart Society (ICISS). IEEE, 2014.

[6] Lukas, Samuel, et al. "Student attendance system in classroom using face recognition technique." 2016 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2016.

[7] Okokpujie, Kennedy O., et al. "Design and implementation of a student attendance system using iris biometric recognition." 2017 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2017.