PURBANCHAL UNIVERSITY

DEPARTMENT OF COMPUTER ENGINEERING
KHWOPA ENGINEERING COLLEGE
LIBALI-8, BHAKTAPUR

A MID-TERM REPORT

ON

Facial Attendance System

A project submitted for the partial fulfillment of requirements for the degree of Bachelor of Engineering in Computer Engineering (Eighth Semester)

SUBMITTED BY:

Amar Nagaju (730305)

Anirudh Basukala (730307)

Manish Kharbuja (730319)

Manish Nhuchhe (730320)

Sunder Tamakhu (730345)

UNDER THE GUIDANCE OF:

Er. Shiva K. Shrestha

SUBMITTED TO:

DEPARTMENT OF COMPUTER ENGINEERING

KHWOPA ENGINEERING COLLEGE

15th March 2021

ABSTRACT

Machine learning is a powerful application performed by powerful computers, and Facial recognition is one of the major applications of machine learning. In the automated attendance system, we use a facial recognition system to record the real-time attendance of recognized faces. In this project, only the paralinguistic features of faces are taken into account to develop a facial attendance system. The project emphasizes frames with human faces to extract the effective features of eyes, nose, mouth within the face and designed reliable classifiers related to different person faces using machine learning and deep learning algorithms. We have used 16 layers of convolution layers with max-pooling for classifying 10 different faces. We have implemented a VGG-16 architecture model, which achieves 99.94% training accuracy and 99.99% validation accuracy in our project. The system in our project records the attendance of staff based on the real-time as they appear in the frame. The interactive Django framework makes it user-friendly and easy to browse and view the record.

Keywords: Facial Recognition, Attendance, feature extraction, CNN

Table of Contents

List of Figures

CHAPTER 1
INTRODUCTION

1.1 Background

The emerging growth of technology and the increasing computational power of computer systems have given rise to the ever-growing dream of Artificial Intelligence to reality. Facial recognition is one of the major feat achieved through machine learning which is a major part of artificial intelligence. Facial recognition is being widely applied in real-life scenarios like security, surveillance, trafficking, and attendance system. Face recognition uses correlation recognition algorithms for face recognition or discrimination, which is based on the extracted face image features[1]. Deep Convolutional Neural Networks (DCNNs) is currently the method of choice both for generative, as well as for discriminative learning in computer vision and machine learning[5].

1.1.1 Face Detection

Before the face recognition stage, we need to detect faces in an image and bound the region defining the face. For this, the software has to know to distinguish between a basic face and the rest of the background. A human face has numerous, distinguishable landmarks, the different peaks, and valleys that make up facial features. It is often defined as nodal points. Each human face has approximately 80 nodal points. Some of the nodal points are:

- Distance between the eyes
- Width of the nose
- Depth of the eye socket
- The shape of the cheekbones
- The length of the jawline

A face print is created with these nodal points in the form of numerical code that represents the face in the database.

1.1.2 Face Recognition

Face Recognition is the process of verifying the identity of a person using their facial landmarks. Facial recognition captures, analyzes, and compares patterns of the person's details of the face. Once face detection is performed, feature extraction is done to provide effective information that is useful for distinguishing between faces of a different person. Facial recognition systems have seen wider uses in recent times on smartphones and other forms of technology. Face recognition is the method of recognizing people through their face images. It has numerous applications in the field of biometrics, security smart card attendance systems, and surveillance systems. The neural network formation is performed through the three most common layers of convolution, activation, and pooling. The convolution layer is responsible for setting an image through convolution layers and prepares a feature map based on it. Activation layer or ReLU helps in quick and efficient model training by activation of

only necessary features node in a neural network. Pooling minimizes the parameters for the network to learn quicker and protects only parameters that aids in learning.

1.1.3 Facial Attendance

In an educational environment, attendance is one of the important factors for employee management and control. The attendance system makes sure of the presence of teachers on the college premises. Deep learning is a fast-evolving subject in the modern world, the system can now automatically detect and keep the track of the presence of the registered bodies as well as record and manage the data. Face recognition is not a new topic in the current world. It has been widely used in the face unlock system in mobile phones. The autofocus system in cameras is very common these days as well. Extending the application of the face detection technique we focus on capturing the same for the attendance purpose in the concerned sector.

The current method of recording the attendance manually is a hectic and time-consuming process. Also, the fingerprint system serves only an employee at a time. Therefore, the automated system works to remove such problems and makes an easier attendance system without having to manually record one after another data.

1.2 Motivation

With the rapid technological advancement, image processing and deep learning are to reach their peak. Utilizing the same opportunity to extend its use we create the system that will make use of the knowledge on same to create the system that can detect, recognize, compare and record the data based on frontal facial recognition. With the automated attendance system, the presence of the concerned person can be known.

1.3 Statement of Problem

In a very short period of machine learning introduction, there has been numerous research conducted in the field of face detection and face recognition. Face recognition has been one of the main applications of machine learning. Yet, the attendance system based on face recognition has barely been explored. There is little research conducted on it but the general usage of the system in the area of schools, colleges, offices are minimum. The systems in use are usually manual or other forms of biometric attendance systems. The manual attendance system is a hectic, slow, and error-prone system. The manual way of taking attendance can easily be tampered with and does not hold much significance in the digital world. Also, other biometric systems are slow in terms that they can record one data at a time. They can be slow and problematic at times.

1.4 Objectives

The main aim of this project is:

- To detect, recognize and record personnel in a user-friendly interface.

1.5 Scope

The major scope of the project are:

- In keeping a record of attendance of staffs in schools and colleges.
- In keeping a record of students in educational and other institutions.
- In surveillance presence of people in college premises.

1.6 Structure of Project Report

This report includes the following chapters:

- Chapter 1 includes the introduction of the project including a background of the study, objectives, problem statement, motivation, and scope of the project
- Chapter 2 presents the literature reviews of the project work.
- Chapter 3 presents the methodology used for the development of the project.
- Chapter 4 shows the task completed for the mid defense of the project work.
- Chapter 5 includes the task to be completed for the final defense of the project work.

CHAPTER 2
LITERATURE REVIEW

Face recognition is a new and growing topic that has taken many researchers by the heart. Many algorithms have been developed in a very short duration of a few years. The earliest work on face recognition can be traced back to the 1950s, the research on automatic machine recognition of faces started in the 1970s but a fully automatic face recognition system based on a neural network was reported back in 1997.

2.1 Face Detection

Face detection has been a topic of intensive attention in recent years. Many of these research provides different perspective like model architecture, data augmentation, label assignment, etc. Tinaface uses ResNet-50 as the backbone to achieve 92.1% average precision (AP) which claims to exceed most of the recent face detectors with a larger backbone[11]. A major face detection technique dlib used facial landmarks in an image like eyes, eyebrows, nose, mouth, jawline applied to face alignment, head pose estimation, face swapping, blink detection, and so on.

2.2 Face Recognition

The dimension of facial features is too large and does not exist with deep learning face recognition, which was a problem[1]. Deep Convolution Neural Networks (DCNNs) have demonstrated impressive results in face recognition[7]. Initially, a formal method of classifying faces was proposed where the author proposed collecting facial profiles as curves, finding their norm. FaceNet showed excellent results and was superior to other methods. By using VGGFace2 pre-trained models, FaceNet can touch 100% accuracy on YALE, JAFFE, AT & T datasets, Essex faces95, Essex grimace, 99.375% for Essex faces94 dataset, and the worst 77.67% for the faces96 dataset. VGG 16 as proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper “Very Deep Convolution Networks for Large-Scale Image Recognition” achievers 92.7% test accuracy in ImageNet. VGG-16 improves AlexNet by replacing large kernel-sized filters with 3×3 sized kernels.

2.3 Transfer Learning

Transfer learning is used to transfer the knowledge gained while solving one problem to solve a similar problem. It can be widely used in the context of object detection like the knowledge gained in recognizing cars can be used to some extent to recognize other vehicles like trucks. We can give the new dataset to fine-tune the pre-trained CNN. It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in a skill that they provide on related problems.

2.3.1 VGG 16

VGG-16, a convolution Network is obtained by training the training dataset, which is used for feature extraction. The VGG-16 is a 16-layer deep convolutional network and uses 3×3 convolution kernels. [1] The input of the net is a fixed RGB image of 224×224; this deep CNN architecture mainly consists of thirteen convolutional layers, five pooling layers, and three fully-connected layers; the last fully-connected layer has 54 channels. ReLU activation functions are used in the convolutional and fully connected layers. It enlarges the training set using the method of data augmentation by generating multiple virtual images from each original image using bilateral filter and image translation. The result from [1] has an accuracy of the accuracy of 86.3%. Its accuracy can be increased to 98.1%. Fine-tuning is applied to improve the efficiency of training and get a better result with fewer iterations [2].

2.4 Facial Attendance

Attendance through card or fingerprint is widely used in many offices and organizations. It is now being replaced with a facial recognition system. The face can be used as an identification for recording attendance daily. The recorded data is used to generate a report and analyze the staff of the organization. For attendance, through Facial Attendance System a person should make his/her face visible to the camera properly. Otherwise, the system may not recognize the person. In this system, there is no need for manual intervention for keeping the record. Face recognition can be an effective way to improve the attendance system. This system can be easily implemented and also saves a lot of time due to its automated working system.

CHAPTER 3
REQUIREMENT ANALYSIS

3.1 Software Requirement

Software requirements for our system include:

a. Anaconda
b. Python
c. Visual Studio Code
d. Keras with Tensor flow backend
e. Matplotlib
f. Skimage
g. PIL
h. Adobe Photoshop
i. Microsoft Word
j. Microsoft PowerPoint
k. Django
l. Flask
m. CUDA
n. GitHub

3.2 Hardware Requirement

Hardware requirements for our prepared system include:

a. GPU-GeForce GTX 1050 Ti (4 GB)
b. RAM (8 GB)
c. Webcam

3.3 Functional Requirement

The functional requirement for the prepared systems are:

a. The system should be able to capture images/videos in real-time.
b. The system should have a login system for users and admin.
c. The system should provide an attendance report to the user.
d. The report must be able to be viewed on basis of the time id of the user.
e. It should have an alert system when data is recorded.
f. It should send an email notification to absentees.

3.4 Non-Functional Requirement

These are not needed by the system but are essential for the better performance of the system. The points below focus on the non-functional requirement of the system prepared.

3.4.1 Reliability

The accuracy of the system is conducted through different test metrics.

3.4.2 Accuracy

The accuracy of the model is defined by the ability to recognize the people as trained datasets successfully. The system should have as maximum accuracy as possible.

3.4.3 Maintainability

Due to the modular nature of each service, backend, frontend, and model development are highly separated. Each can be independently maintained.

3.4.4 Portability

As all the processing is done on the server, the web application can be viewed from any browser.

3.4.5 Performance

The face recognition requires less processing power in comparison to the model development. So the system can provide better performance.

3.4.6 Frames per Second (FPS)

The face detection and recognition need to be done precisely in as many frames of a video as possible in a short time for reliability.

CHAPTER 4

FEASIBILITY STUDY

Various estimation is made of whether the identified user needs may be satisfied using current software and hardware technologies. The following points describe the feasibility of the project.

4.1 Technical Feasibility

The project requires lots of labeled images for each individual added as part of the recognizable faces. Collecting images for each individual is the biggest challenge to the project. The images need to be labeled in individual classes for each person which is further divided into training, validation, and testing labels. The dataset can be collected with a python script. The script can detect a face in real-time and recorded video and crop and save the image. Training the large dataset requires large computational power. VGG algorithm can be used as a transfer learning algorithm to pass on some of the basic features to the trained model that makes the training easier, accurate, and require fewer resources. So, our project is technically feasible.

4.2 Operational Feasibility

The project is based on the first thing performed in the schools and colleges i.e. checking if the students and staffs are present in the college premises. It also holds applications in the field of offices and other organizations. The project can be used in such areas to reduce time-wasting, create applicable reports. It also provides an interactive and user-friendly interface that makes it easier to view the records. So, our project is operationally feasible.

4.3 Economic Feasibility

The main expenditure of the project lies in computation. For computation, we used a high-end PC with GPU availability. We used GeForce GTX 1050 Ti for model training owned by one of the team members and also used our laptops for other tasks. For capturing images for the dataset we used a webcam connected to the laptop. So, our project is economically feasible.

CHAPTER 5

SYSTEM DESIGN AND ARCHITECTURE

The system design and architecture can be featured in the following sections:

5.1 Architecture of Face Recognition System

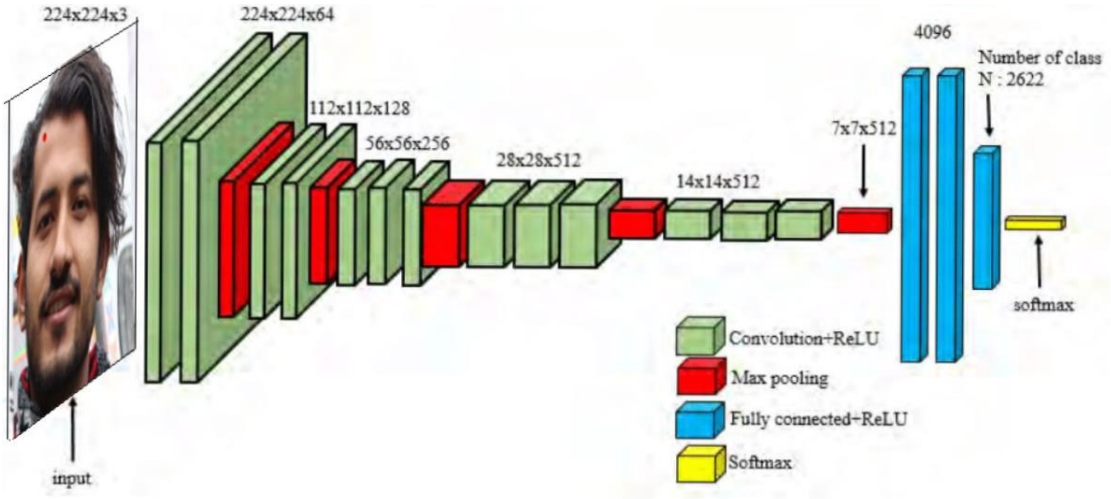

Figure 5.1: Architecture of Face Recognition System

We used VGG-16 architecture to implement our project.VGG-16 is convolutional neural network architecture; its name VGG-16 came from the fact that it has 16 layers. Its layers consist of Convolutional layers, Max Pooling layers, Activation layers, Full connected layers. There are 13 Convolutional layers, 5 Max Pooling layers, and 3 Dense layers which sum up to 21 layers but only 16 weight layers. Conv 1 has several filters as 64 while Conv 2 has 128 filters, Conv 3 has 256 filters while Conv 4 and Conv 5 has 512 filters.

The most unique thing about VGG-16 is that instead of having a large number of hyper-parameter they focused on having convolution layers of 3*3 filters with a stride 1 and always used the same padding and max pool layer of 2*2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end, it has 2 FC (fully connected layers) followed by a softmax for output. The 16 in VGG-16 refers to it has 16 layers that have weights.

5.2 System Design

System design is the process of defining the elements of a system such as the architecture, modules, and components, the different interfaces of those components, and the data that goes through the system. It is meant to satisfy the specific needs and requirements of a business or organization through the engineering of a coherent and well-running system.

5.2.1 Block Diagram

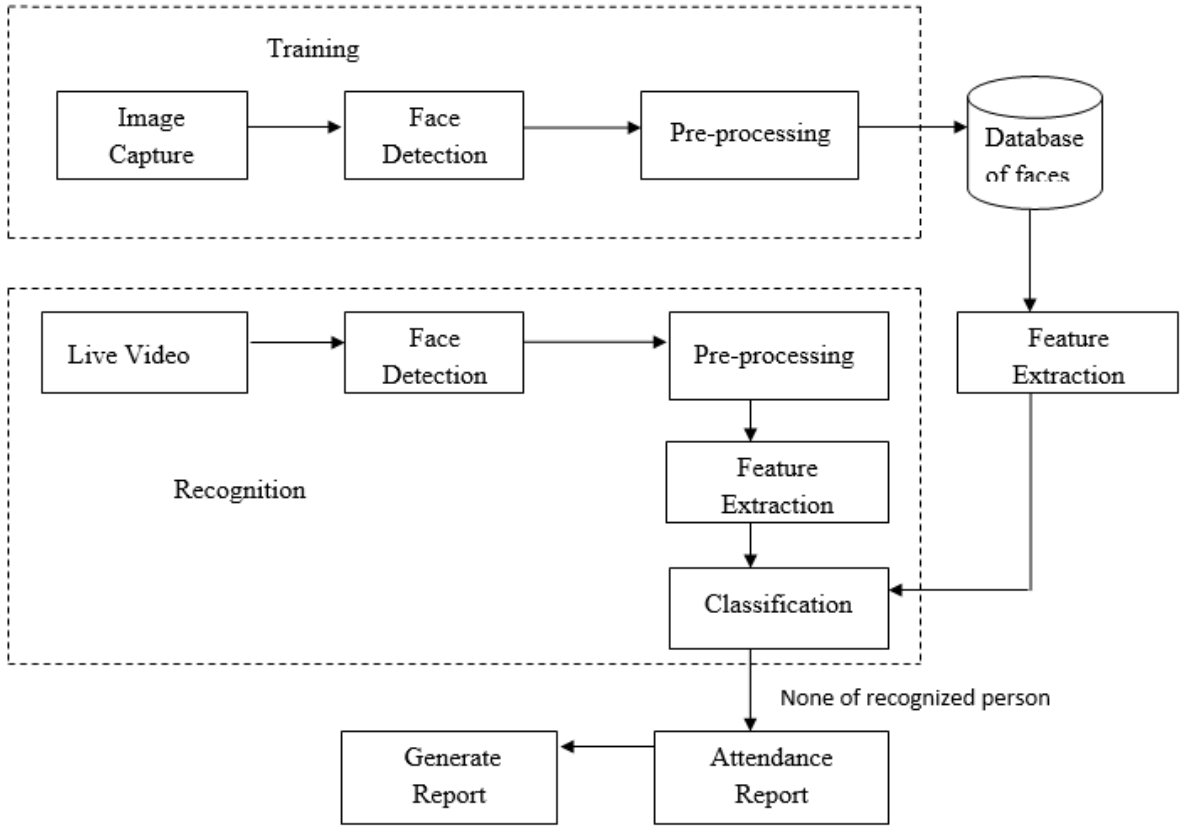

Figure 5.2: Block Diagram of Facial Recognition Attendance System

An environment is everything in the world in which an agent is present. Here physical agents are agents situated in a physical real world. In our system, the person is an agent in the physical environment. Therefore, image acquisition of a person is done in our system, to extract the features of a person i.e. from the face. Where we can define image acquisition as the creation of a digitally encoded representation of visual characteristics of an object, such as a physical scene or the interior structure of an object. Likewise, our system will also train the image capture with preprocessing of data for making dataset with features extraction. And AI is defined as the process of a machine imitating human intelligence factors such as; learning, self-correction, and reasoning. So, training data is often used in the validation and testing set of data. These two extracted features that are from the image acquisition and training dataset are classified according to their classification. Classification in machine learning is when a machine or computer uses an algorithm to conclude data that it already has, and then uses these conclusions to categories new data it receives. Hence, image recognition is done through the help of classification. Finally recognizing the person, our system automatically records the attendance of the person in the database.

5.2.2 Use-Case Diagram

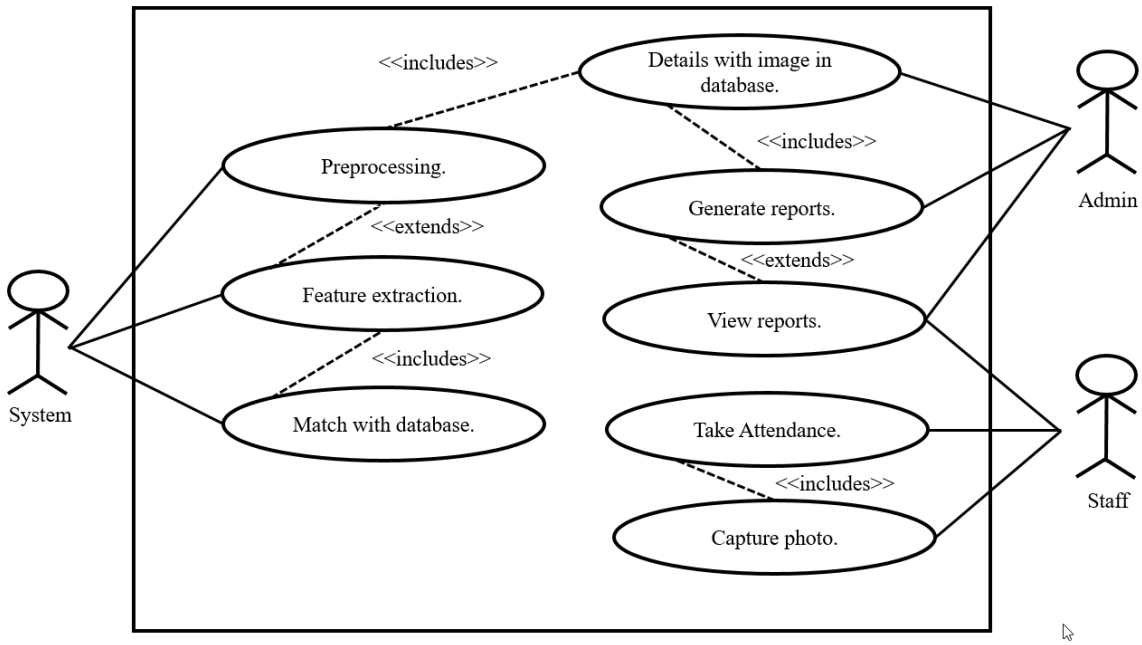

Figure 5.3: Use-Case Diagram

For our system, we used three actors in the use-case diagram. Capturing the photo of the Student (i.e. first actor) is done. The system actor is responsible for preprocessing the captured photo and extracted features of the photo as per the requirement. Also, matches the extracted features with the database and records the attendance report of those students by the system itself. The final actor i.e. Admin is an overall manager who manages all the details of the image of a student in the database.

5.2.3 ER-Diagram

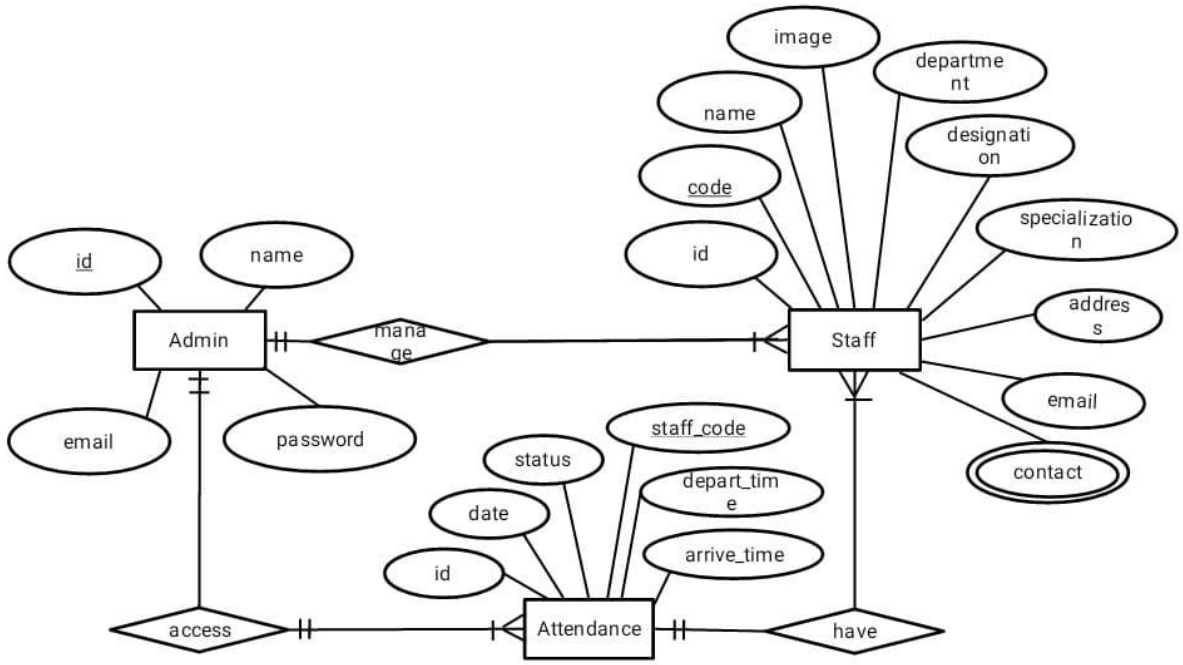

Figure 5.4: ER diagram of Facial Attendance System

As a high-level conceptual data model diagram, we have constructed the above ER-diagram for Facial Attendance System. Admin, Staff and Attendance are the main entity of this ER-diagram. This ER-diagram of the Facial Attendance System explains the logical structure of an entity stored in a database. Here entity of Admin consists of attributes like id as primary key, name, email, and password. Entity Staff consists of attribute like code as primary key, name, image, department, designation, specialization, contact, email, address. Likewise, the Attendance entity consists of the attribute of code, Status, arrive_time, depart_time, date, and id. For the interaction of entities, we have a relationship between the above entities. These relationships can be described as one and only one Admin entity can manage one or more Staff Entity. Then Staff Attendance Entity records attendance with Time in and Time out. This Attendance record can access by admin and can also be seen by staff by login into the web application.

5.2.4 Context Diagram

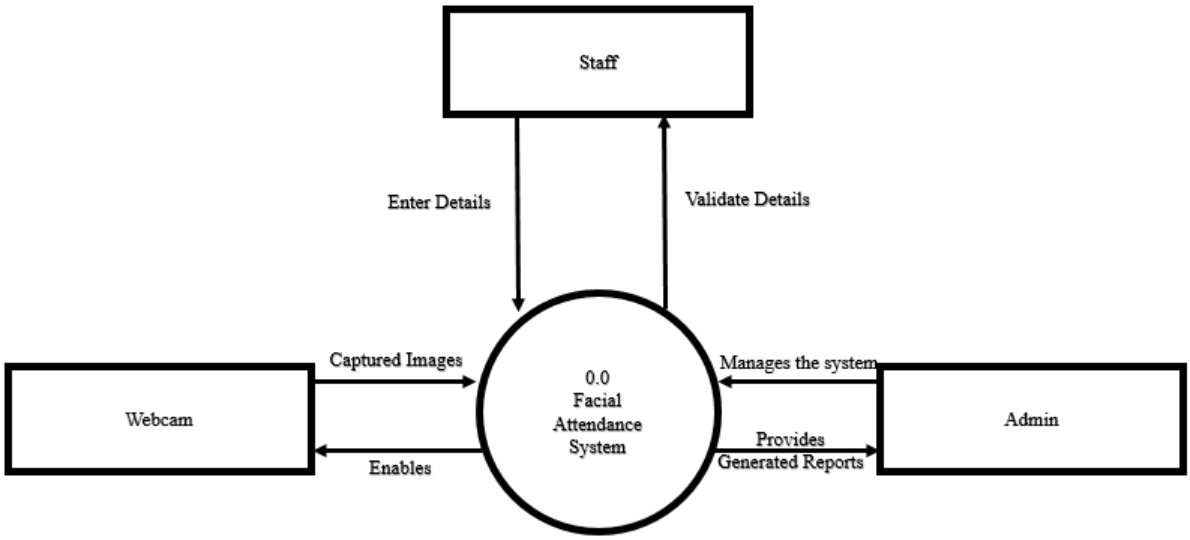

Figure 5.5: Context Diagram

The diagram in figure 6 is the context diagram of the "Facial Attendance System". It explains the basic entities involved in the system and the basic operation performed in the system. According to the figure, the system shows the relationships of three entities: Staff, Admin, and Webcam with the Facial Attendance System. A webcam is responsible for capturing the data of the faces that are fed into the system. The model we created, recognizes the personnel and records the attendance. Staff is the ones whose image is to be recognized. Admin uses the Django framework to control the system.

5.2.6 Data Flow Diagram

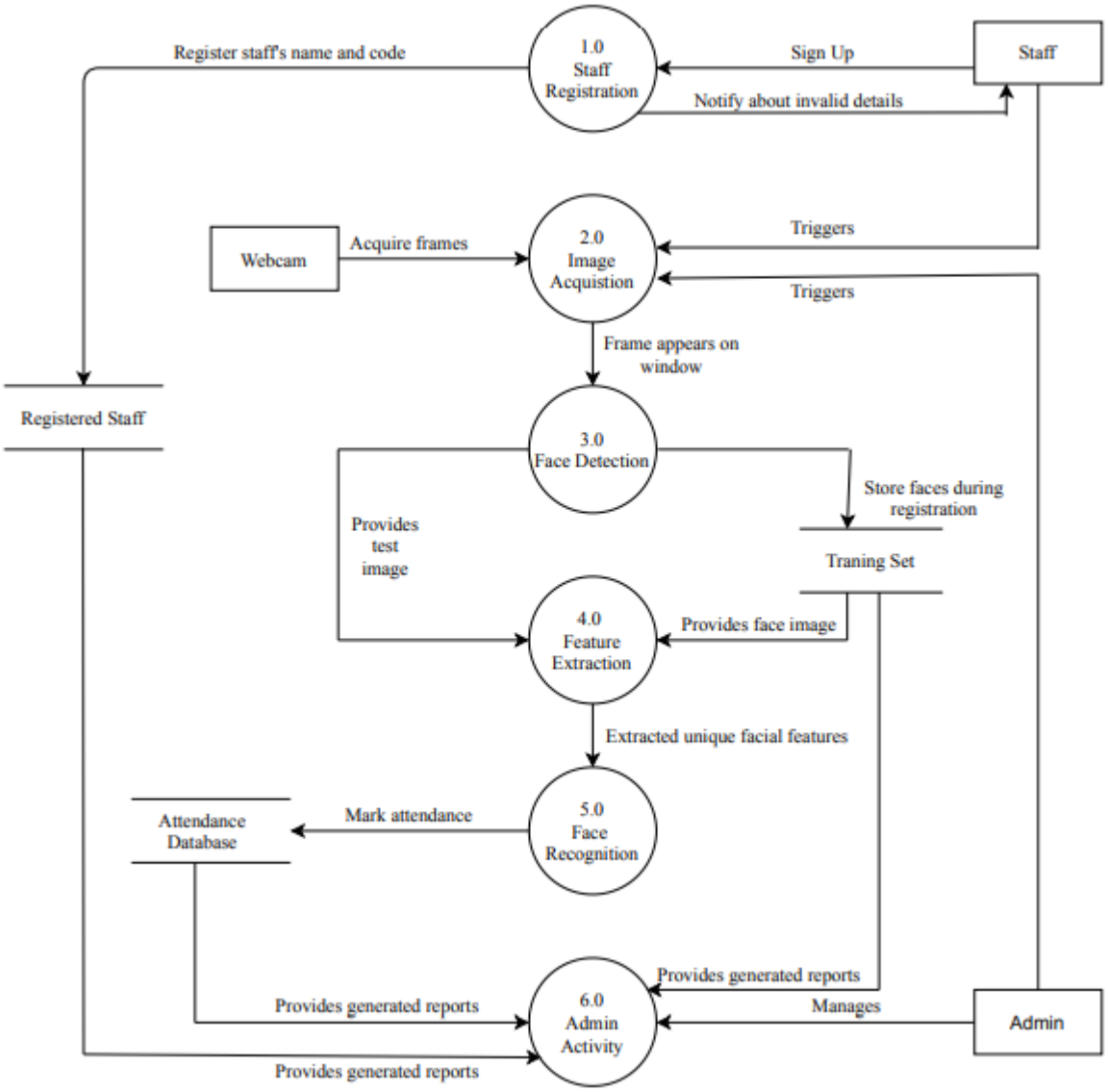

Figure 5.6: Data Flow Diagram Level 1

The data-flow diagram of level 1 for our system can be seen in the above figure. This DFD level 1 is an "exploded view" of the context diagram of the Facial Attendance System. Here in this diagram, sub-processes of the context diagram, that is Staff registration, Image Acquisition, Face Detection, Features Extraction, Face recognition, and Admin activity can be seen in the above figure. Beside Registered staff, the Attendance database and Training Set are databases used in this system.

In this DFD level-1 figure, Staffs are signed up using the staff registration sub-system and these data are saved in the registered staff database, which can be accessed by the Admin Activity sub-system. Likewise, Staff's image and image frame are acquired by using a webcam. These frames are either stored for the training set database or there are used for the testing image. Then features are extracted by the Feature extraction sub-system and unique features are provided to the Face recognition sub-system which recognizes the person and

marked their attendance in Attendance Database. These databases are accessed by the Admin activity sub-system for generating reports and managing these reports.

5.2.5 Sequence diagram

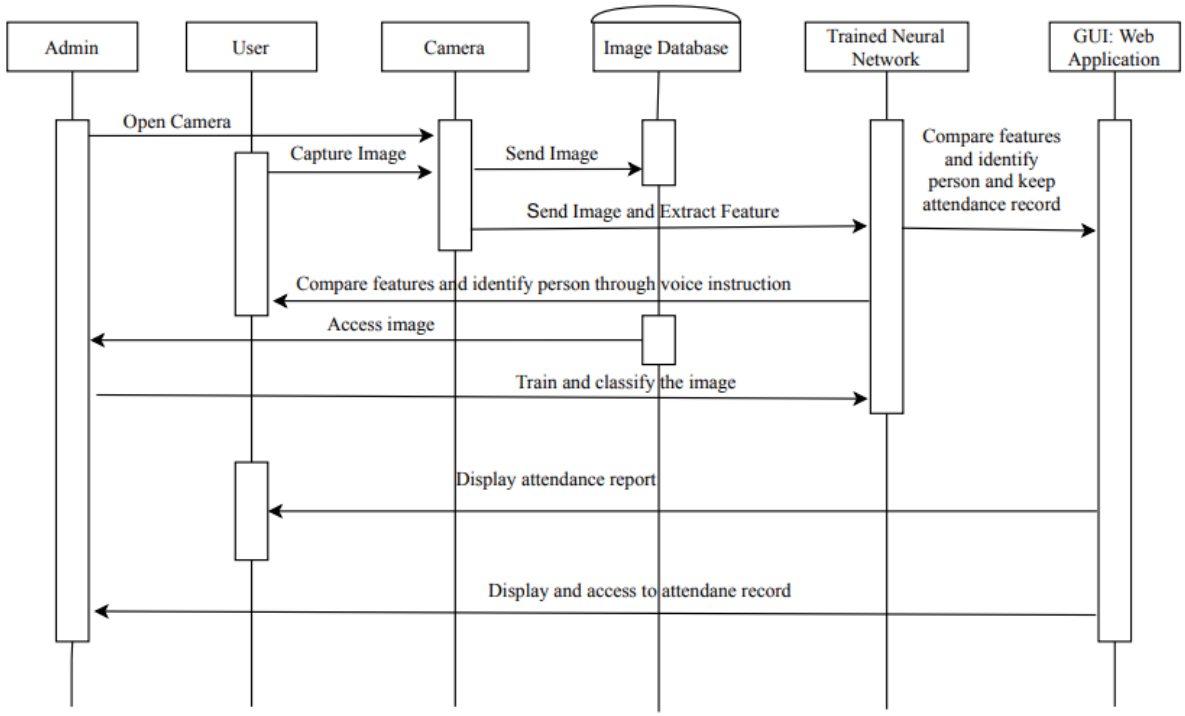

Figure 5.7: Sequence diagram of Facial Attendance System

The diagram in figure 7 is the sequence diagram of the “Facial Attendance System”. According to the figure above, the admin looks after the system and when the admin starts the application, the camera is accessed. The image obtained is sent to the image database and the model is created through a trained neural network. The record-keeping is performed and can be viewed through the web application. In the meantime, the trained model compares features and when a person is identified it sends an alert to the user. Through the image database, further images can be collected and trained through transfer learning to add more personnel to the system. The web-based interface is then able to display the attendance to the user and also to the admin.

CHAPTER 6
METHODOLOGY

6.1 Background

For developing the Facial Recognition System, various phases and methods are used to proceed with the help of various software, tools, and languages. In our project, we have gathered data in form of an image and use it to train the system and later utilize it to recognize the person. The architecture used for the project is specified further in the report.

6.2 Deep Learning

Deep learning is a method by which a computer is trained to perform the task as performed by humans, such as speech recognition, classifying images, making predictions, human face-recognizing. Such a task is assisted by the neurons-like structure called a neural network that learns and passes on the information it learned for further learning. CNN-based learning is a widely used and most popular learning approach that is quick and more accurate than any other technique.

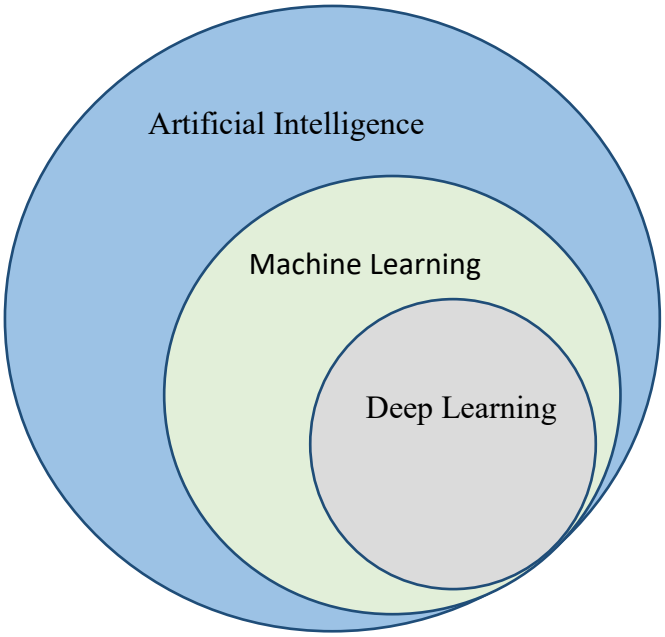

Figure 6.1: Deep Learning in AI

6.3 Dataset Selection

6.3.1 Dataset 1

The dataset was prepared manually for 10 people in this project work. The dataset consists of 7500 images. The dataset consists of 10 classes of a person with each containing 750 images per person. Out of which 525 images per person were used for training the model while 150 images per person were used for validation and the rest 75 images for testing. So total number of images for training was 5250 images for 10 people whereas the total number of images for validating was 1500 images for 10 people and 750 images were used for testing.

This was done to maintain the ratio of 7:2:1 or simply 70% for training purposes, 20% for validation, and 10% for testing purposes. The dataset was created by running the model at 10 epochs and 20 epoch. Also, we found 20 epochs brought more accuracy and less loss than the 10-epoch model.

Figure 6.2: Dataset created for model creation

6.3.1.1 Experimental Analysis

Models with different approaches were trained. The dataset was manipulated constantly with the hope of a better result each time. Different augmentation techniques like rotation, scaling, resizing, etc. were performed. We used learning rate optimization techniques to further increase the model accuracy. The X-axis shows the change in epochs and Y-axis shows the accuracy and loss.

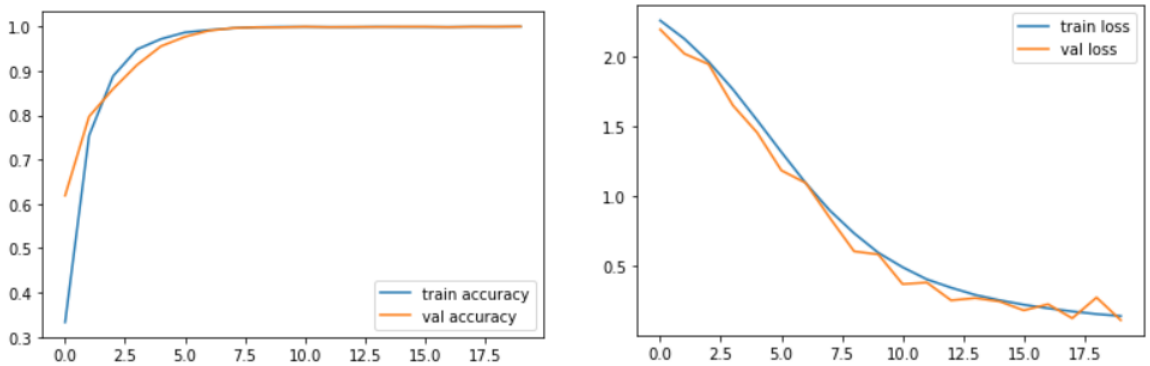

Figure 6.3: Accuracy and Loss graph of dataset 1

6.3.2 Dataset 2

The second dataset we created consists of 45 classes. In this, we used a single image to create the face recognition model. The single image was manipulated into creating multiple images. We used a package of python called albumentation to manipulate the image. Albumentation is a python library used for fast and flexible image augmentations. It makes quick performance for producing a wide variety of image transform operations for performance. It performs such augmentation while providing a concise, yet powerful image augmentation interface for different computer vision tasks.

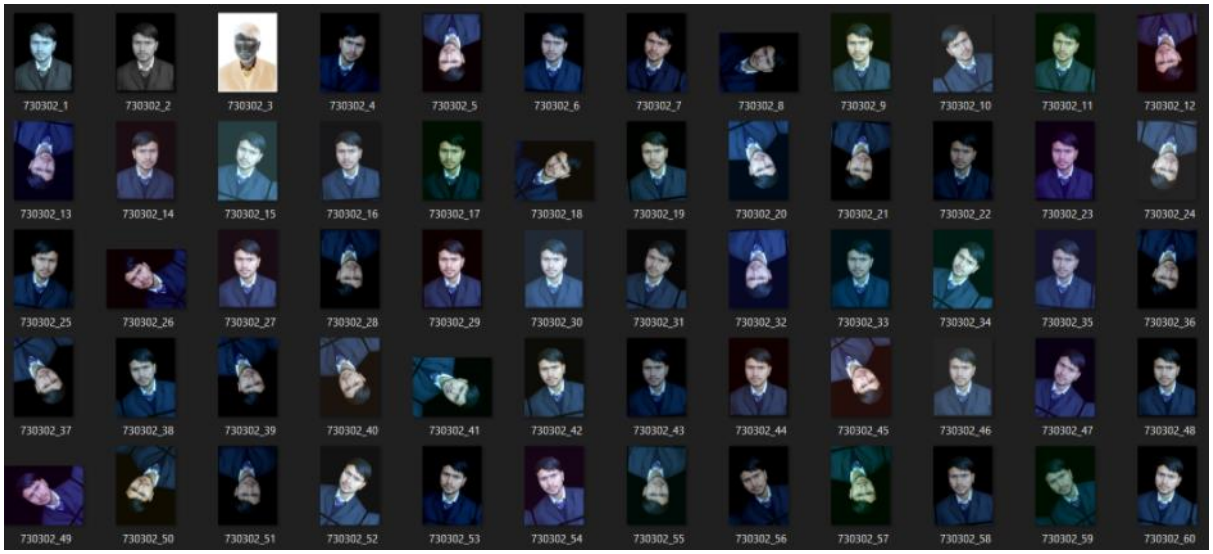

Figure 6.4: Image created with albumentation

Various albumentation techniques are:

a. Rotation
b. Flip
c. Transpose
d. Motion blur
e. Median blur
f. Blur
g. Distortion
h. Brightness
i. Saturation
j. Hue
k. Contrast

Although the albumentation provided with various augmentation techniques, the dataset created was discarded as it was understood that using 1 image for model creation was not applicable as the accuracy of the model dropped drastically and the real-time test was a complete failure. Dataset 2 was set to be a fail and didn't provide any significant benefit to the model.

6.3.2.1 Experimental Analysis

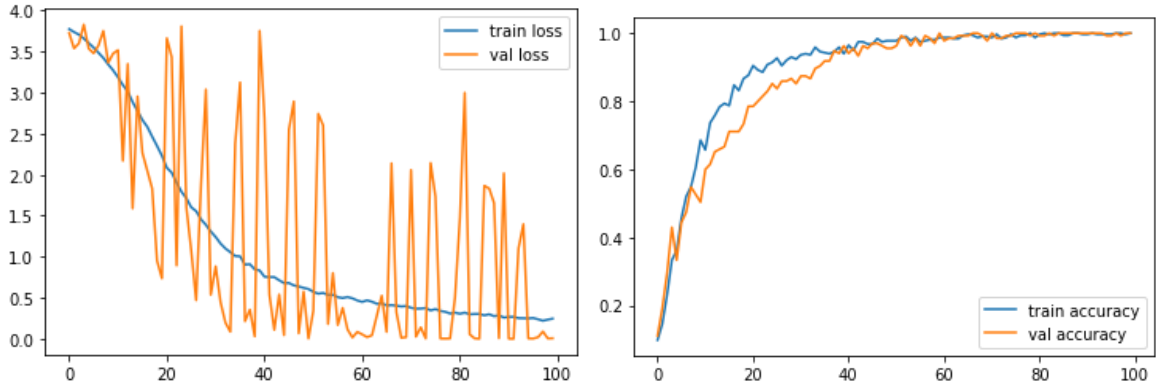

Figure 6.5: Accuracy and Loss graph of Dataset 2

The model given in the above figure shows the training graph of 45 faces. We have used only one image per face where these images were augmented by using the albumentation library. The Albumentations library was used for image augmentation and different pixel-level transform of the flip, random brightness, RGB shift, gray, invert image was used in this augmentation to create 10 images per face. There is a total of 450 images of 46 faces where we have used seven images per face for the training dataset and three images per face for the validation dataset. Then we trained this model running 100 epochs. As a result, training accuracy was 0.33, and validation accuracy was 0.61. Likewise, training loss accuracy was 2.2541, and validation loss was 2.19. Here we can see the training model achieves overfitting while training. Moreover, this model barely recognizes about 4-5 faces only and other faces are predicted randomly.

6.4 Dataset Map

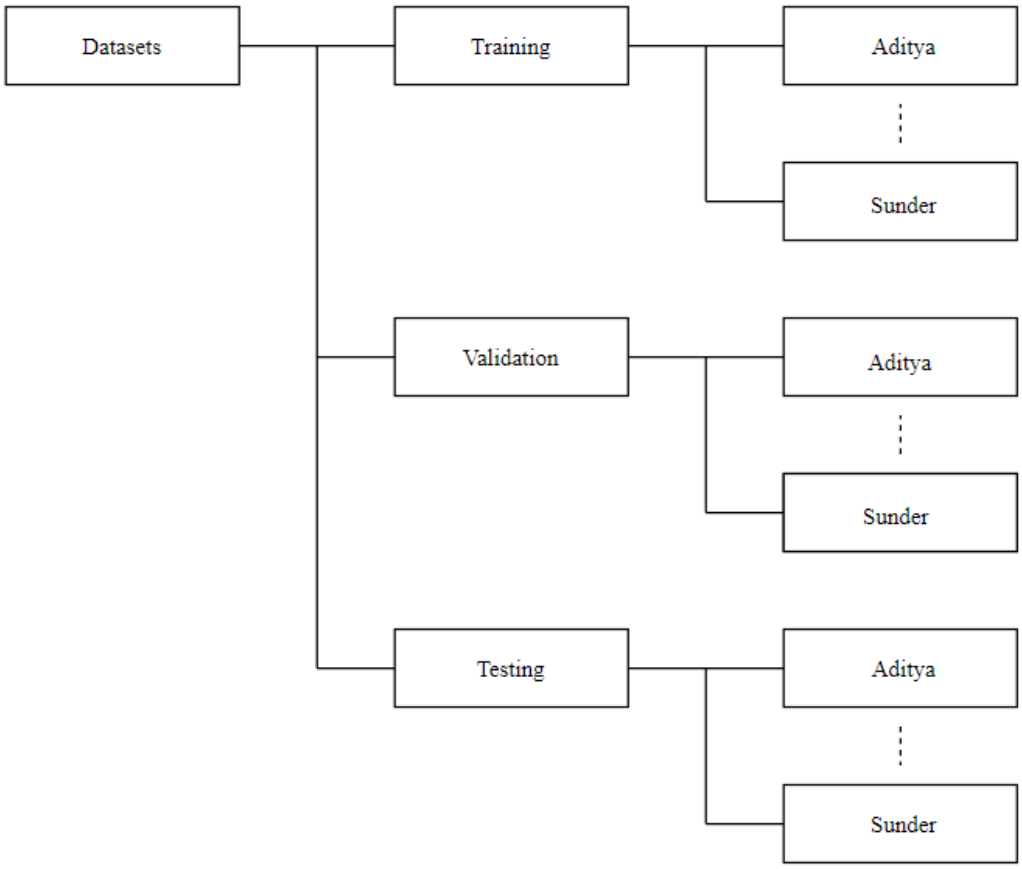

Figure 6.6: Mapping of the dataset

The dataset is divided into 3 sets. The training dataset is used by the model to train the model. It contains 70% of the images. The validation dataset is used by the model to check the accuracy of the trained model. It contains 20% of the dataset. The testing dataset contains 10% of the images and is used to test the accuracy of the model.

6.5 Data Augmentation

We used different data augmentation techniques during data preprocessing to create the variation in the dataset. The following techniques were applied to each image.

1. Rotation
2. Scaling
3. Rescaling
4. Zoom
5. Horizontal Flip

6.6 Parameter Selection

To prepare a model for face recognition, various parameters need to be defined. The selection of these parameters aids in the proper learning of the input images by the computer. These parameters are defined before training the dataset. Based on the obtained output the necessary

changes are made to the parameters like changing the number of images in the dataset, size of the image, no. of hidden layers, learning layers, and so on. The following are the parameters selected for this project work.

- Image size :224*224
- Batch size :8
- No. of epochs :10
- Learning rate : 0.0001
- Activation function :ReLU

CHAPTER 7
WORK DONE

7.1 Overview

With half the time due for the project, several of the tasks have been completed with some still in progress. The tasks like face detection with the recognition of 10 personnel have been completed. For the easy interaction between the user and the system, a Django framework makes the viewing of the records much easier. The database has been created and successfully connected with the system so the record can be kept safely for recording data and keeping it for safekeeping for further use.

7.1.1 Face Recognition

A model has been successfully created that can successfully recognize 10 people. The dataset was trained with 20 epochs. We have collected 7500 facial pictures with dimensions of 224*224.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
conv1_1 (Conv2D)	(None, 224, 224, 64)	1792
conv1_2 (Conv2D)	(None, 224, 224, 64)	36928
pool1 (MaxPooling2D)	(None, 112, 112, 64)	0
conv2_1 (Conv2D)	(None, 112, 112, 128)	73856
conv2_2 (Conv2D)	(None, 112, 112, 128)	147584
pool2 (MaxPooling2D)	(None, 56, 56, 128)	0
conv3_1 (Conv2D)	(None, 56, 56, 256)	295168
conv3_2 (Conv2D)	(None, 56, 56, 256)	590080
conv3_3 (Conv2D)	(None, 56, 56, 256)	590080
pool3 (MaxPooling2D)	(None, 28, 28, 256)	0
conv4_1 (Conv2D)	(None, 28, 28, 512)	1180160
conv4_2 (Conv2D)	(None, 28, 28, 512)	2359808
conv4_3 (Conv2D)	(None, 28, 28, 512)	2359808
pool4 (MaxPooling2D)	(None, 14, 14, 512)	0
conv5_1 (Conv2D)	(None, 14, 14, 512)	2359808
conv5_2 (Conv2D)	(None, 14, 14, 512)	2359808
conv5_3 (Conv2D)	(None, 14, 14, 512)	2359808
pool5 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc6 (Dense)	(None, 4096)	102764544
fc6/relu (Activation)	(None, 4096)	0
fc7 (Dense)	(None, 4096)	16781312
fc7/relu (Activation)	(None, 4096)	0
fc8 (Dense)	(None, 2622)	10742334
fc8/softmax (Activation)	(None, 2622)	0

Total params: 145,002,878
Trainable params: 145,002,878
Non-trainable params: 0

Figure 7.1: Model Architecture of Face Recognition Model

We designed a CNN model to solve this problem. The face detected with the face detection network is passed to the input of this model and the model processes the image to classify to the respective classes to make a model ready for recognition. Our model used VGG-16 as the base model and added our layers on the top of the VGG-16 layer.

7.1.2 Web Framework

Django is a widely-used Python web application framework. Django is used as it develops websites fast, clean, and with practical design. We created user-friendly interactive GIS web applications that harness the power of spatially enabled data. The reports of absentees along with the present staff can be viewed easily with the click of a button. The basic outlay of the framework has been created which can be easily manipulated as per requirements.

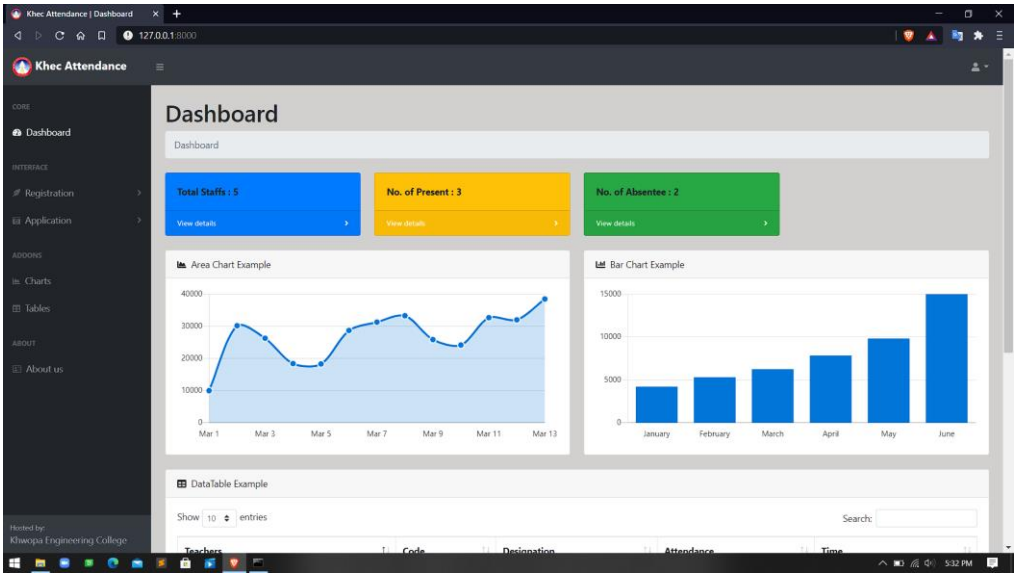

Figure 7.2: Web framework dashboard

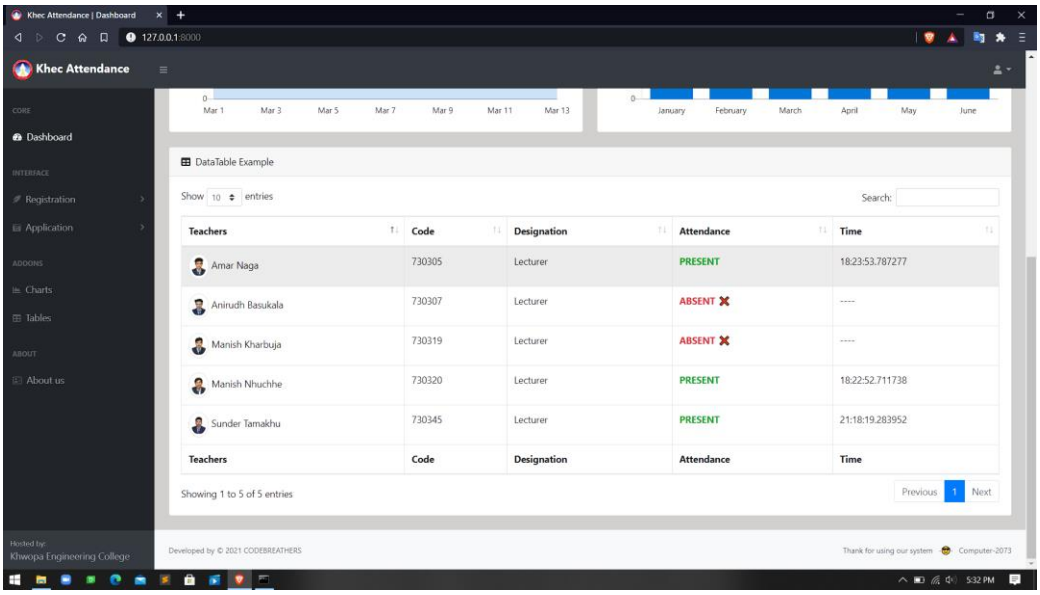

Figure 7.3: Viewing the data in the Web framework

4.1.3 Data Dictionary

For our project, we used the MySql database service to create the database. We created tables for storing information about staff and recording attendance. A daily record of attendance is stored in the database, which is later retrieved to generate reports. Various attributes are used in the table to keep the information.

Attributes of Staff information:

1. name
2. code (primary key)
3. department
4. designation
5. specialization
6. email
7. contact
8. address

Attributes for Attendance record:

1. id (primary key)
2. s_code (foreign key)
3. date
4. time
5. status

Besides these tables, other tables are created to store the username, password for user login, and admin login. Username and password are given to each user to login and see the report of their daily attendance record.

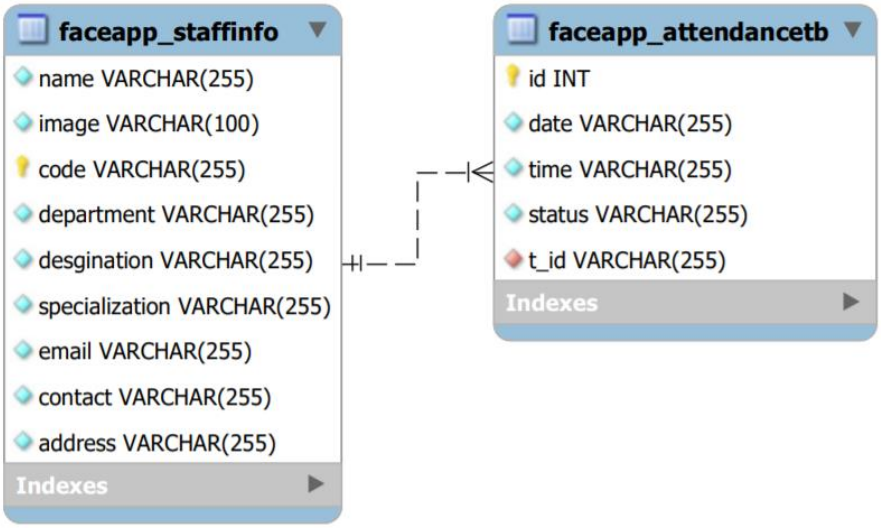

Figure 7.4: Data dictionary of Facial Attendance System

4.3 Test Result

The model for face recognition was created with the augmentation of the dataset during pre-processing. The augmentation like rotation, flipping of the images, scaling is done to vary the dataset for better accuracy in the model in real-time execution. The training accuracy of the model was found to be 0.98.

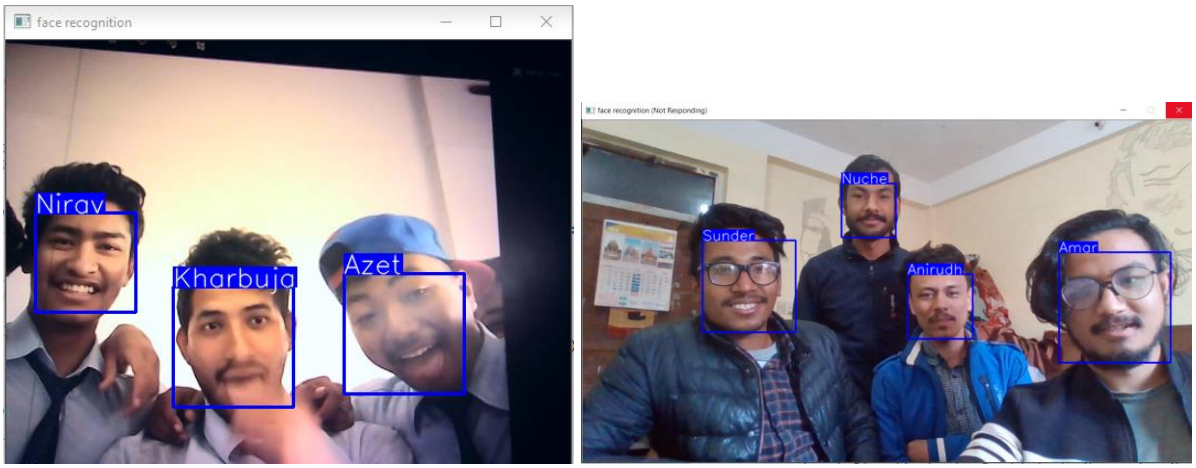

Figure 7.5: Test Result for face recognition

4.3.1 Failed Cases

Despite the correct predictions, initially, there were few failed predictions by the model.

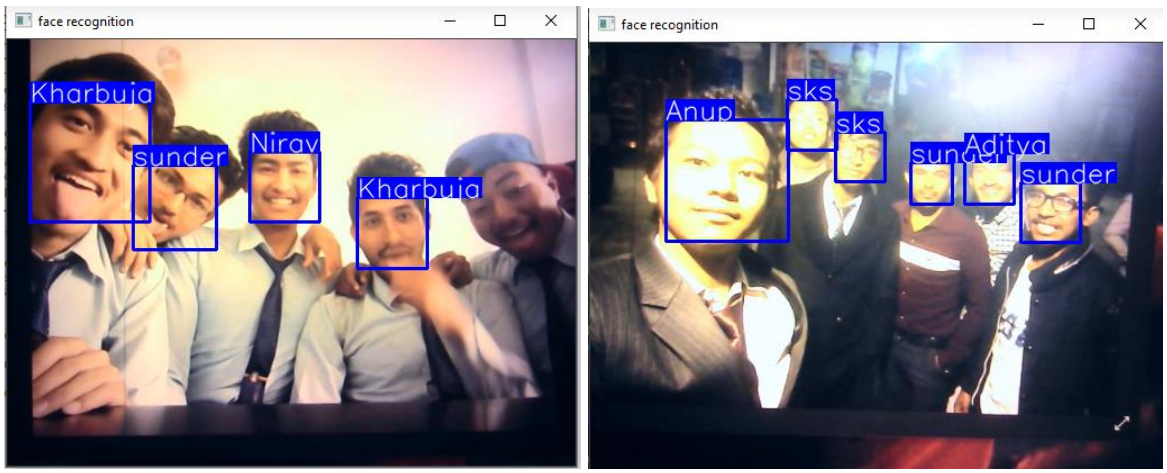

Figure 14: Failed prediction of the face recognition model

The failed cases were seen due to lack of training dataset, a dataset of a person with high-quality frame result increase in accuracy rate and lack of augmentation of dataset decrease the accuracy rate. Besides, lighting also played a major factor in the wrong prediction of the detected faces by the face recognition model. Low lighting predicate the person randomly as sufficient light is not present in his/her face to detect the person. Eventually camera quality uses to detect the face also affects the recognition. Using a high-quality camera can easily detect the faces easily and recognize the person by our system whereas a dull or low-quality camera barely detects the face and recognize the person. Hence these factors lead to failure in the test case.

CHAPTER 5

WORK TO BE DONE

Despite the work that we mentioned in the previous section, some tasks are yet to be done. The remaining work for the project completion are:

1. Login for a user to see their attendance report
2. Admin page completion
3. Train model with more dataset
4. Add alert when personnel is recognized and recorded
5. Email notification for absentees

REFERENCE

[1] Hongling Chen and Chen Haoyou, “Face Recognition Algorithm Based on VGG Network Model and SVM”, IOP Conf. Series: Journal of Physics: Conf. Series 1229 (2019) 012015 IOP Publishing doi:10.1088/1742-6596/1229/1/012015, [Accessed: 28thFebruary, 2020]

[2] Zhao Pei, Hang Xu, YanningZhang, Min Guo and Yee-Hong Yang, “Face Recognition via Deep Learning Using Data Augmentation Based on Orthogonal Experiments”, 25 September 2019. [Visited Date: 28th February 2020]

[3] SrikanthTammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images", Electrical Engineering, Indian Institute of Technology, Hyderabad, 2017 [Accessed: 28th February 2020]

[4] Mei-Xiang Zhao, Zhi-Gang Jia, Dunwei Gong "Sample-Relaxed Two-Dimensional Color Principal Component Analysis for Face Recognition and Image Reconstruction", [Accessed: 28th February 2020]

[5] MengjiaYan Horizon Robotics, Qian Zhang Horizon Robotics, Guoli Wang Horizon Robotics, ZiningXu Horizon Robotics "VarGFaceNet: An Efficient Variable Group Convolutional Neural Networkfor Lightweight Face Recognition", [Accessed: 3rd March 2020]

[6] Felipe L. Gewers, Gustavo R. Ferreira, Henrique F. de Arruda, Filipi N. Silva, Cesar H. Comin, Diego R. Amancio, and Luciano da F. Costa1 "Principal Component Analysis: A Natural Approach to Data Exploration”, [Accessed: 10th February 2020]

[7] Grigorios G. Chrysos, Stanlianos Moschoglou, Giorgos Bousitsas, Jiankang Deng, Yannis Panagakis, Stefanos Zafeiriou, "Deep Polynomial Neural Network", 2021

[8] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, “Retinaface: Single-stage dense face localization in the wild “, 2020

[9] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “Arcface: Additive angular margin loss for deep face recognition,” in Conference on Computer Vision and Pattern Recognition (CVPR), 2019

[10] G. Chrysos, S. Moschoglou, G. Bouritsas, Y. Panagakis, J. Deng, and S. Zafeiriou, “π´nets: Deep polynomial neural networks,” in Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[11] Yanji Zhu, Hongxiang Cai, Shujan Zhang, Chenhao Wang, Yichao Xiong, “Tinaface: Strong but Simple Baseline for face Detection”, 2020

APPENDIX

A.1 Test case of Face Recognition System

Figure A.1: Correct prediction by the model

A.2 Recorded data in the framework

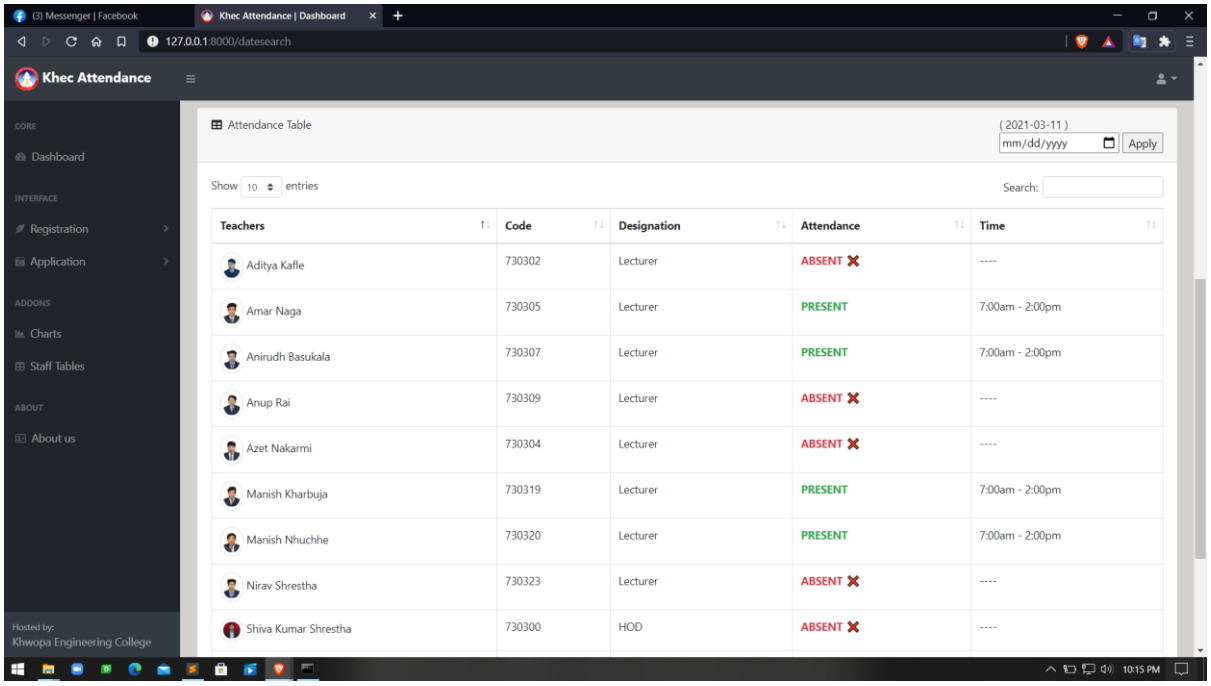

Figure A.2: Recorded data after correct prediction

A.3 Viewing data based on date

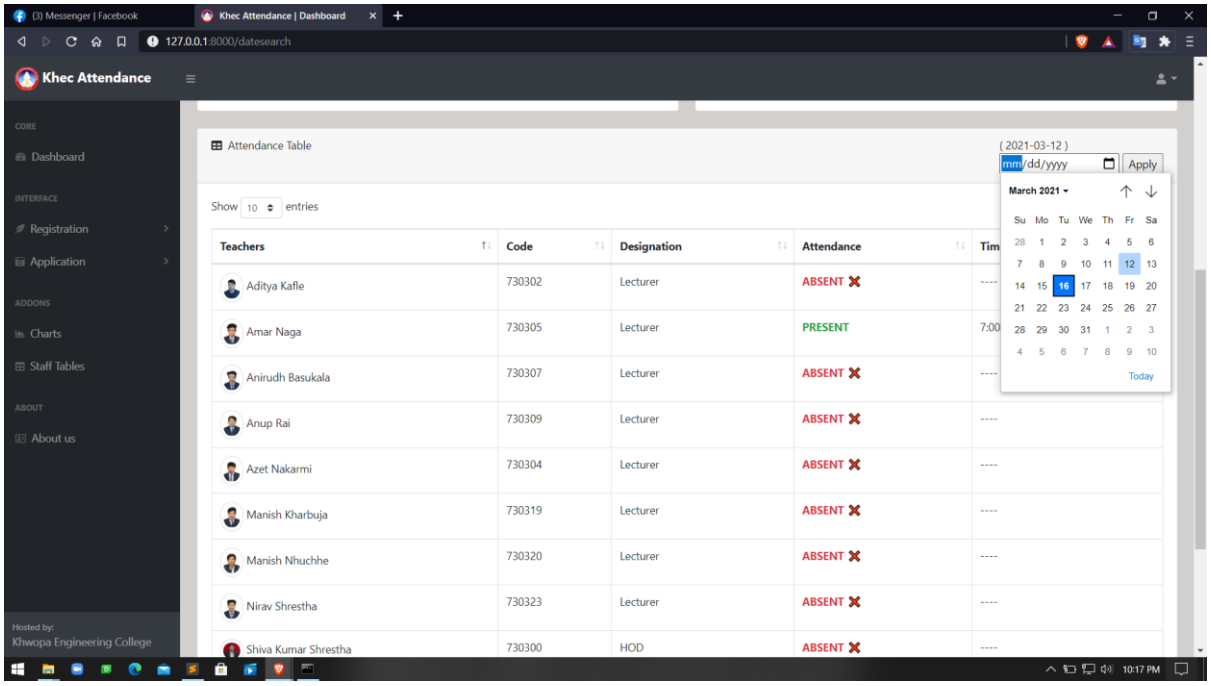

Figure A.3: Viewing data on specific date

A.4 Failed case with model trained with dataset 2

Figure A.4: Failed prediction by the model