**PURBANCHAL UNIVERSITY**

****

**DEPARTMENT OF COMPUTER ENGINEERING**

**KHWOPA ENGINEERING COLLEGE  
LIBALI-8, BHAKTAPUR**

**A MID-TERM REPORT**

**ON**

**"Face Recognition System"**

A project submitted for the partial fulfillment of requirements for the degree of Bachelor of Engineering in Computer Engineering (Seventh 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

13 March 2020

# ABSTRACT

This system that we have build briefly describes about the seventh semester project on "Facial Recognition System”. Normally, facial system is collected and managed manually through the use of deep learning. This project attempts to solve the task of low reliability of other provided biometric artificial intelligence and use facial recognition system by analyzing patterns with the features of textures, shape of the persons face. This system can detect and successfully recognize five people as per the dataset used to train the model. The use of popular neural network makes the system to train and learn at high rate and aid in successful facial recognition. It has great learning ability which uses various features for creating a model which can be used to determine the required output. Providing that there are many algorithms to recognize the face we used VGG-16 approach which helped to achieve a better result in recognizing the face. With the dataset we created of total 6250 images the system provides an accuracy measure of 99%.

Keywords: *Biometric AI, Facial Recognition, Convolution Neural Network, VGG-16*

Contents

[ABSTRACT ii](#_Toc34993555)

[List of figures iv](#_Toc34993556)

[List of Abbreviations v](#_Toc34993557)

[INTRODUCTION 1](#_Toc34993558)

[1.1 Background 1](#_Toc34993559)

[1.2 Motivation 2](#_Toc34993560)

[1.3 Statement of Problem 2](#_Toc34993561)

[1.4 Objectives 2](#_Toc34993562)

[1.5 Our Approach 2](#_Toc34993563)

[1.6 Scope 2](#_Toc34993564)

[1.7 Structure of Project Report 2](#_Toc34993565)

[LITERATURE REVIEW 3](#_Toc34993566)

[METHODOLOGY 5](#_Toc34993567)

[3.1 Background 5](#_Toc34993568)

[3.2 Deep Learning 5](#_Toc34993569)

[3.3 Block Diagram 6](#_Toc34993570)

[3.4 Algorithm used 7](#_Toc34993571)

[3.5 Parameter Selection 8](#_Toc34993572)

[3.6 Data Set 8](#_Toc34993573)

[3.7 Tools and Platform 8](#_Toc34993574)

[WORK DONE 10](#_Toc34993575)

[4.1 Overview 10](#_Toc34993576)

[4.2 Test Analysis 10](#_Toc34993577)

[4.3 Experimental Analysis 12](#_Toc34993578)

[WORK TO BE DONE 14](#_Toc34993579)

[Reference 15](#_Toc34993580)

[APPENDIX A 16](#_Toc34993581)

# List of figures

|  |  |  |
| --- | --- | --- |
| 1.1 | Structure of a CNN | 1 |
| 3.1 | Deep Learning in AI | 5 |
| 3.4 | Block Diagram of Face Recognition System | 6 |
| 3.3 | VGG-16 neural network architecture | 7 |
| 4.1 | Error in prediction | 10 |
| 4.2 | Test Result | 11 |
| 4.3 | Wrong Prediction | 11 |
| 4.4 | Loss and Accuracy of model1 | 12 |
| 4.5 | Loss and Accuracy of model2 | 12 |
| 4.6 | Loss and Accuracy of model3 | 13 |
| A.1 | Test image for 1 person | 16 |
| A.2 | Test image for 5 people | 16 |

# List of Abbreviations

|  |  |
| --- | --- |
| CNN | Convolution Neural Network |
| VGG | Visual Geometry Group |
| GPU | Graphical Processing Unit |
| ReLU | Rectified Linear Unit |
|  |  |
|  |  |
|  |  |

CHAPTER 1

# INTRODUCTION

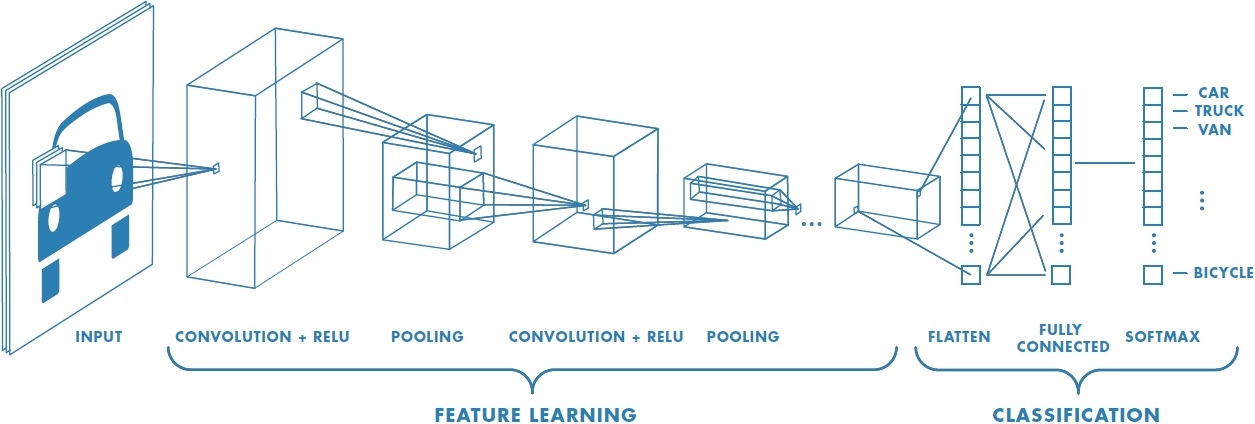
### Background

Face Recognition is the method of recognizing people through their face images, with numerous applications in the field of biometrics AI, security, smart cards, attendance systems and surveillance systems. [8] Face recognition is one of the most popular and growing biometric technologies. With the development of technology, the accuracy and effectiveness of face recognition has increased. Now there are many algorithms and approaches which are being used in face recognition. These all algorithms use different features to develop a model.

In this system, the process starts by the input device detecting the frontal face of the person matching it with the faces in the database and based on which the person is matched to the available images and hence record the available data.

Face recognition is not really anew topic in the current world. It has been widely used in the face unlock system in the mobile phones. The auto focus system in cameras is very common these days as well. Extending the application of the face detection technique we focus in capturing the same for the attendance purpose in the concerned sector.

Convolution networks (ConvNets) have dealt a great deal of success in in large-scale image and video recognition due to availability of large-scale image datasets and high-endGPUs. [7]After the emergence of deep learning, the idea of face recognition has drastically changed. The CNN (Convolution Neural Network) is one of the popular neural networks being used for various purposes. It has great learning ability which uses various features for creating a model which can be used to determine the required output. The neural network formation is performed through three most common layers of convolution, activation and pooling. The convolution layer is responsible for setting an image through convolution layers and prepare a feature map based on it. Activation layer or ReLU helps in quick and efficient model training by activation only necessary features node in neural network. Pooling minimizes the parameters for the network to learn quicker and protects only parameters that aids in learning.



*Fig. 1.1: Structure of a CNN*

*Source: https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html*

VGG-16 is a deep convolution network for object recognition developed and trained by Oxford’s renowned ‘Visual Geometry Group’, which achieved very good performance on the ImageNet dataset.

### 1.2 Motivation

With the rapid technological advancement, image processing and deep learning is to reach its peak. Utilizing the same opportunity to extend its use we wish to 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 the frontal facial recognition.

### 1.3 Statement of Problem

1. The traditional methods of biometric technologies are slow and time consuming.
2. Other biometric approaches are less reliable.

### 1.4 Objectives

The main aim of the project is to create a system that will detect and recognize the face of certain people.

### 1.5 Our Approach

Our approach is to prepare a model for facial recognition using CNN and VGG-16 architecture.

### 1.6 Scope

* Attendance System
* Surveillance System
* Smart Cards
* Security

### 1.7 Structure of Project Report

This report includes following chapters:

* Chapter 1 includes the introduction of the project including 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 2

# LITERATURE REVIEW

In [1], here they had stated that their project is based on VGG and SVM network. Dimension of facial features is too large and does not exist with deep learning face recognition which was a problem for this project. Beside this project extract face features, reduce features dimensions and avoid irrelevant feature to participate in calculation. VGG-16, a convolution Network is obtained by training the training dataset, which is used for feature extraction, and PCA (Principal Component Analysis) is used for feature dimensionality reduction and lastly face recognition is performed by SVM (Support Vector Machine) classifier with linear kernel function. An SVM algorithm is used to predict sample, the basic idea of SVM is to map the input vector into a high dimensional space by non-linear transformation.

The authors Hongling Chen and Chen Haoyou initiate their paper by describing the need of biometric recognition and its advancement in security monitoring and human interaction. Here in this project, the extraction method is based on geometric features of face and treats face image as random vector which include Eigen face, independent component analysis and singular value decomposition. The VGG-16 used in this paper is a 16-layer deep convolutional network and use 3\*3 convolution kernels. The authors have performed experiment on two datasets. The LFW (Label Face in the Wild) dataset is used to evaluate the algorithm in this experiment as well as CelebA (CelebFaces Attributes Datasets). LFW, a dataset composed of more than 13000 pictures concerning worldwide celebrities through the Internet, contains more 5000 people. Each face image has its unique ID and serial number to distinguish it. Celeb A is a large face attribute dataset with more than 200 thousand face images and each image has 40 attribute annotations. In the experiment, they have compared their method and VGG-16, and got highest accuracy in 400 dimensions. Beside they also compare state-of-the-art face recognition method on LFW dataset and got 97.47% result of accuracy.

In [2], VGG-16 network pre-trained with a VGG- Face dataset is used for face recognition. 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 accuracy of accuracy of 86.3%, which is higher than PCA. 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.

In [3], it uses one of the pre-trained models- VGG-16 with Deep Convolutional Neural Network to classify images. Transfer of knowledge is a method of reusing pre-trained model knowledge for another task. VGG-16 architecture consists of 13 convolutional layers, 2 fully connected layers and 1 Softmax classifier. Input image (RGB image with depth 3) is passed into first and second convolutional layer and dimension is changed to 224\*224\*64. The third and fourth convolutional layers are followed by max pooling layer. The fifth, sixth and seventh layers use 256 feature maps. The eight to thirteen layers have 512 kernel filters. The fourteen and fifteen layers are fully connected hidden layers followed by a softmax output layer.

The first model build using convolutional neural network gives validation accuracy of 72.40 %. Then with fine tuning, this model with image augmentation and achieved accuracy of 79.20 %. Eventually with use of one of the pre-trained models (VGG-16) trained on huge dataset of images and fine-tuned with image augmentation to achieve accuracy of 95.40 %

According to [4], to improve the discriminative and generalization ability of lightweight network for face recognition, they propose an efficient variable group convolutional network called VarGFaceNet. This network solves conflict between small computational cost and unbalance of computational intensity inside a block. In order to enhance the discriminative ability of VarGNet for large scale face recognition task, they first add SE block and PReLU on blocks of VarGNet. Then we remove the down sample process at the start of network to preserve the more information. For setting of VarGNet, SE block is adding on normal block of VarGNet which is the down sampling block. Beside it does not use downs ample in first convolution in order to keep enough information. For embedding setting of VarGNet, channel isexpanded from 320 to 1024. Then variable group convolution is employed and point wise convolutions is used to reduce the parameters and computational cost, while remain essential information.

According to [5], 2-Dimensional Princpal Component Analysis(2DCPA) extract the spatial information and best features of 2D samples which can improve the performance of dimensional reduction. 2DPCA is to find a sub space (called Eigen faces or features). The 3 aspect of 2DCPA are abstracting the features of matrix samples in both row and column direction, being innovatively armed with generalized ability and weighting the main components by corresponding Eigen values. Beside it used Linear Discriminate Analysis (LDA) which is another powerful feature extraction algorithm and it maximize the ratio of between class distance to with class distance. R2DPCA utilizes the labels (if known) of training sample and can enhance the total scatter of whole project sample. And R2DPCA also generate weighted covariance matrix in newly proposed approach of face recognition. The objective of 2DPCA is to find left and/or right orthogonal bases vectors so that projected matrix samples have largest scatter after projection. 2DPCA is an unsupervised method and overlooks the potential or known label information of samples. 2DPCA avoid the over fitting and to enhance the generalization ability. In large amount of experiments, R2DPCA sufficiently utilizes the label (if known) of training samples, and can enhance the total scatter of whole projected samples. R2DPCA never suffers from the small samples size (3S) problem as supervised method such as LDA. R2DPCA 2 parts are weighting vector and objective function relaxation.

CHAPTER 3

# METHODOLOGY

### 3.1 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 image and use it to train the system and later utilize it to recognize the person. We have used the VGG-16 architecture to train our dataset to prepare our face recognition model.

### 3.2 Deep Learning

Deep learning is a method by which a computer is trained to perform task as performed by humans, such as speech recognition, classifying images, making predictions, human face recognizing. Such task is assisted by the neurons like structure called 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 techniques.

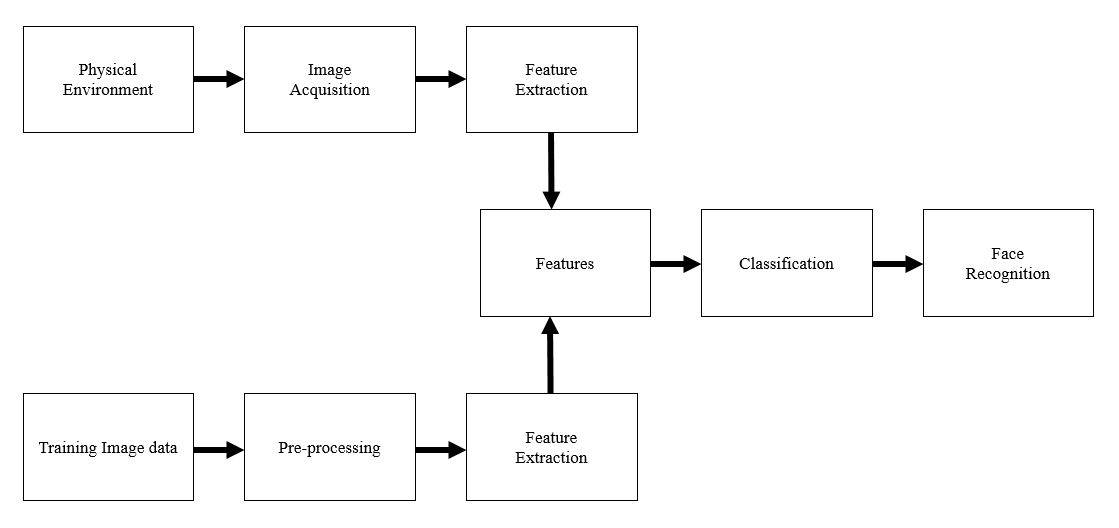
Artificial Intelligence

Machine Learning

Deep Learning

*Fig. 3.1: Deep Learning in AI*

### 3.3 Block Diagram



*Fig. 3.2: Block Diagram of Face Recognition System*

An environment is everything in the world in which an agent is present. Here physical agents is agents situated in a physical real world. In our system individual person is agent in physical environment. So image acquisition of person is done in our system, in order 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 feature that is from 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 draw conclusion from data that it already has, and then uses these conclusions to categories new data it receives. Hence finally the image recognition is done through the help of classification.

### 3.4 Algorithm used

Algorithm

Step1: Start

Step 2: Capture all persons’ face images;

Step 3: Augment training samples and establish the training dataset;

Step 4: for each iteration i to max iterations do

initialize learning rate;

Step5: for each training image I do

Step6: Compute the output of each layer by the forward propagation algorithm

end for

Step7: Update weight ω by the back-propagation algorithm

end for

Step8: for each test image do

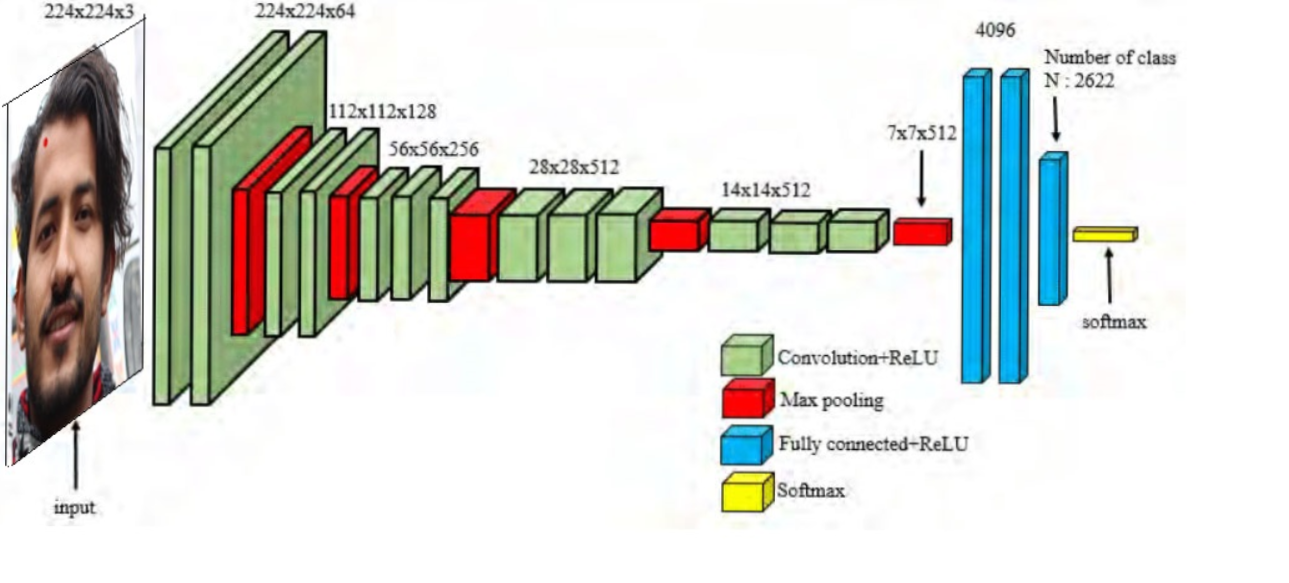
Compute the output of each layer using forward propagation algorithm

Compute the output of network;

Print the person’s name;

end for

Step 9:Stop



*Fig. 3.3: VGG-16-neural-network-architecture*

We used VGG-16 architecture to implement our project.

VGG-16 is convolutional neural network architecture; its name VGG-16 came from 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 sums up to 21 layers but only 16 weight layers. Conv 1 has number of filters as 64 while Conv 2 has 128 filters, Conv 3 has 256 filters while Conv 4 and Conv 5 has 512 filters.

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 same padding and maxpool 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.

### 3.5 Parameter Selection

To prepare a model for face recognition, various parameters need to be defined. The selection of these parameters’ aids in 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 dataset, size of 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

### 3.6 Data Set

The dataset was prepared manually for five of the members of this project work. The dataset consists of 6250 images. The dataset consists of 5 classes with each containing 1250 images. Out of which 1000 images were used for training the model while 250 were used for validation. This was done to maintain the ratio of 8:2 or simply 80% for training purpose and 20 % for validation.

### 3.7Tools and Platform

1. Python Programming Language
2. Anaconda
3. Keras with TensorFlow backend
4. Python Packages

* Matplotlib – allows to create graphs
* Skimage – image manipulation
* PIL – adds support for opening, manipulating, and saving many different image file format

1. Os Environment – Windows
2. Hardware

* GPU-GeForce GTX 1050 Ti (4 GB)
* RAM (8 GB)

1. Design Tools

* Adobe Photoshop

1. Documentation Tools

* Microsoft Word
* Microsoft Powerpoint

CHAPTER 4

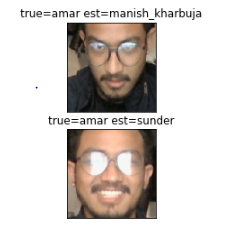
# WORK DONE

### 4.1 Overview

Our project(Face Recognition system) has successfully detected person face as per training data model. Beside we had trained 5-person dataset using 10 epochs with accuracy 0f 0.9994. We have collected 6250 facial pictures of dimension of 255\*255 in order to create a dataset for our facial system. Among those facial pictures 5000 picture were used for training our system where as remaining 1250 picture were used for validation purpose. And the time taken for train those data with 10epochs was about 14.6667min with 140ms/step for each epoch. In order to train our model, we have used GeForce GTX 1050 graphics card with intel i7 8th generation processor.

### 4.2 Test Analysis

Three models were created for the same dataset with some preprocessing in order to check the change in accuracy of model. First the dataset was trained with horizontal flip , shear range of 0.2 and zoom range of 0.2. The training accuracy of model was found to be 0.98 but the test result showed otherwise. There was flickering of predicted name and to remove the error next model was created.



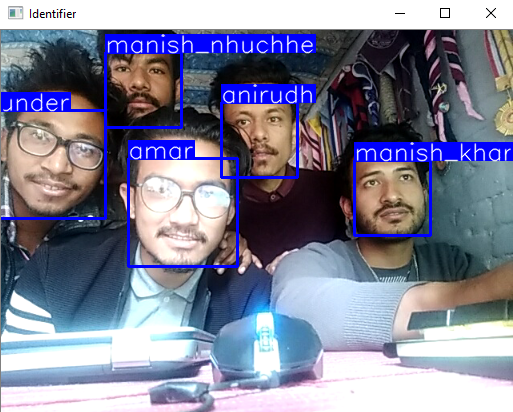
*Fig. 4.1: Error in prediction*

The dataset was preprocessed again with additional rotation of 90 degree and vertical flip. This was done only to the validation set and not on training set. The validation error was unstable averaging 1.2 while the validation accuracy dropped to 0.4.

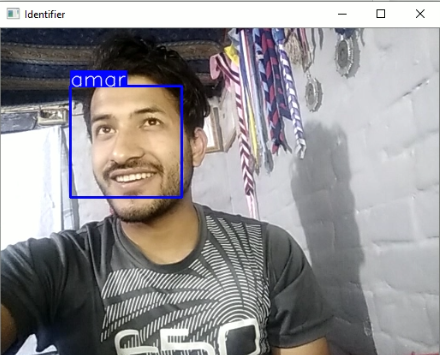
The dataset was once again processed with rotation of 90 degree and additional vertical flip. This time the training and validation accuracy both increased to 0.8.

**4.2.1 Test Result**

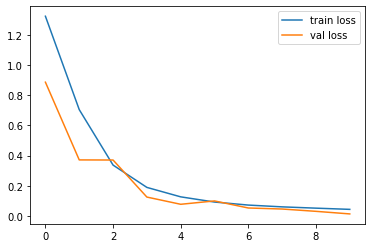
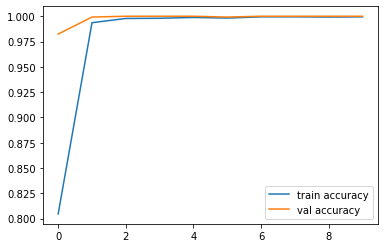
It was found that random flipping and rotating image before model training increases the accuracy of learning.

*Fig. 4.2: Test Result*

**4.2.2 Failed Case**

****

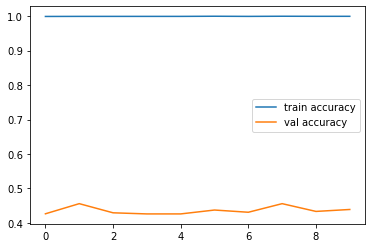
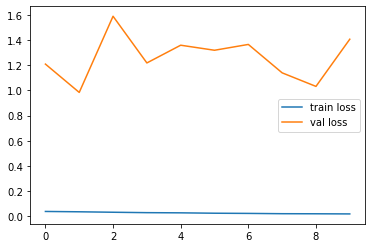
*Fig. 4.3: Wrong Prediction*

4.3 Experimental Analysis  


*Fig. (a) Fig. (b)*

*Fig. 4.4: Loss and Accuracy of model1*

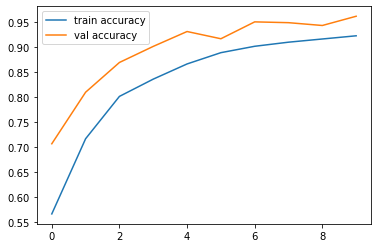
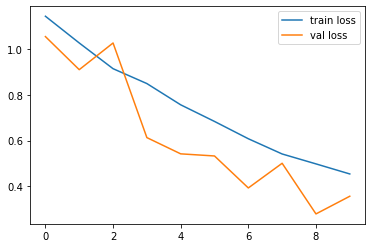
This above graphs shows the losses and accuracy for our first training model.The training loss and validation loss are getting close to each other after the epochs goes by. But in accuracy graph training accuracy curve increased at first 2 epochs and then remains steady along with validation accuracy.



*Fig. (c) Fig. (d)*

*Fig. 4.5: Loss and accuracy of model2*

In this graphs of model2, we can see huge gap between the curves in both accuracy and loss graph. This means there is case of overfitting in this model. The validation loss is much greater than training loss which signifies there is a case of overfitting in model2.



*Fig. (e) Fig. (f)*

*Fig.4.6: Loss and Accuracy of model3*

This graph shows that model3 is best among the 3 models. There is comparatively less gap between the curves in graph. In loss graph, we can see the values decreasing for both curves whereas in accuracy graph there is increase in value for both curves.

CHAPTER 5

# WORK TO BE DONE

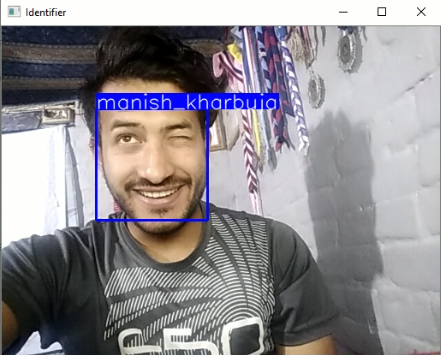
Even though we had trained a set of data and process an output, the result was not exactly what we were hoping for. And some of works left to be done are

* We came to know that our training images were not enough to detect the unknown person.
* We have captured all the photo with same intensity of light while preparing the data set. Now we need to capture photo with different intensity and in different environment background for new data set.
* Beside our system does not count the number of people detected in the frame.
* Lastly a better GUI is left to be created for interaction with the user.

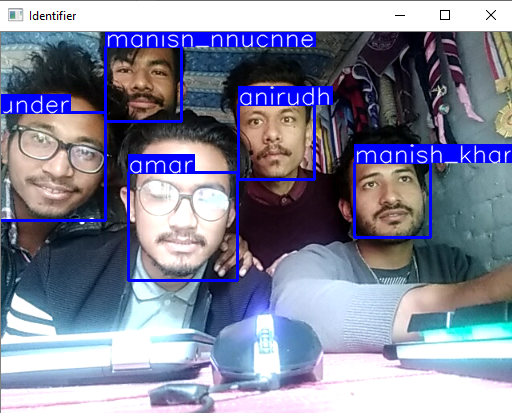
# Reference

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: 28thFebruary, 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: 28thFebruary, 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. Karen Simonyan& Andrew Zisserman, “*VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION”*, Visual Geometry Group, Department of Engineering Science, University of Oxford [Accessed: 28th February 2020]
8. Dr. PriyaGuptaa, NidhiSaxenaa, MeetikaSharmaa, JagritiTripathia, “*Deep Neural Network for Human Face Recognition*”,Maharaja Agrasen College, University of Delhi [Accessed: 3rd March 2020]

# APPENDIX A

**

*Fig. A.1: Test image for 1 person*



*Fig. A.2: Test image for 5 people*