A PROJECT REPORT ON

"IMPLEMENT HUMAN FACE RECOGNITION"

SUBMITTED TO THE UNIVERSITY OF PUNE, PUNE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

BACHELOR OF ENGINEERING
Computer Engineering

BY

SAMARTH MALI	**ROLL NO:41146**
SARTHAK NIRGUDE	**ROLL NO:41153**
KEDAR PAWAR	**ROLL NO:41158**

UNDER THE GUIDANCE OF

PROF. S.N. GIRME

DEPARTMENT OF COMPUTER ENGINEERING
PUNE INSTITUTE OF COMPUTER TECHNOLOGY, PUNE

2024-25

C E R T I F I C A T E

This is to certify that the Project Report entitled

"Patient Readmission Prediction"

Submitted by

SAMARTH MALI	**ROLL NO:41146**
SARTHAK NIRGUDE	**ROLL NO:41153**
KEDAR PAWAR	**ROLL NO:41158**

is a bonafide work carried out under the supervision of Prof. S. N. GIRME and it is submitted towards the partial fulfillment of the requirement of Savitribai Phule Pune University, Pune for the award of the degree of Bachelor of Engineering(Computer Engineering).

Prof. S. N. GIRME
(Guide)

Dr. Geetanjali Kale
(Head of Department
of Computer Engineering)

Dr. S.T. Gandhe
(Principal,Pune Institute of Computer
Technology, Pune)

Internal Examiner

External Examiner

Place: Pune

Date:

ABSTRACT

Face recognition is a rapidly developing and widely applied aspect of biometric technologies. Its applications are broad, ranging from law enforcement to consumer applications, and industry efficiency and monitoring solutions. The recent advent of affordable, powerful GPUs and the creation of huge face databases has drawn research focus primarily on the development of increasingly deep neural networks designed for all aspects of face recognition tasks, ranging from detection and preprocessing to feature representation and classification in verification and identification solutions. However, despite these improvements, real-time, accurate face recognition is still a challenge, primarily due to the high computational cost associated with the use of Deep Convolutions Neural Networks (DCNN), and the need to balance accuracy requirements with time and resource constraints. Other significant issues affecting face recognition relate to occlusion, illumination and pose invariance, which causes a notable decline in accuracy in both traditional handcrafted solutions and deep neural networks. This survey will provide a critical analysis and comparison of modern state of the art methodologies, their benefits, and their limitations. It provides a comprehensive coverage of both deep and shallow solutions, as they stand today, and highlight areas requiring future development and improvement. This review is aimed at facilitating research into novel approaches, and further development of current methodologies by scientists and engineers, whilst imparting an informative and analytical perspective on currently available solutions to end users in industry, government and consumer context.

Contents

Chapter 1

SYNOPSIS

1.1 Project Title

Implement Human Face Recognition

1.2 Technical Keywords

Human Face Recognition, Computer Vision, Image Processing, Facial Features Extraction, Pattern Recognition, Machine Learning, Deep Learning, Convolutional Neural Networks (CNN)

1.3 Problem definition

We Build and Implement Human Face Recognition.

Chapter 2

TECHNICAL KEYWORDS

2.1 Technical Keywords

Human Face Recognition, Computer Vision, Image Processing, Facial Features Extraction, Pattern Recognition, Machine Learning, Deep Learning, Convolutional Neural Networks (CNN)

2.2 Area of Project

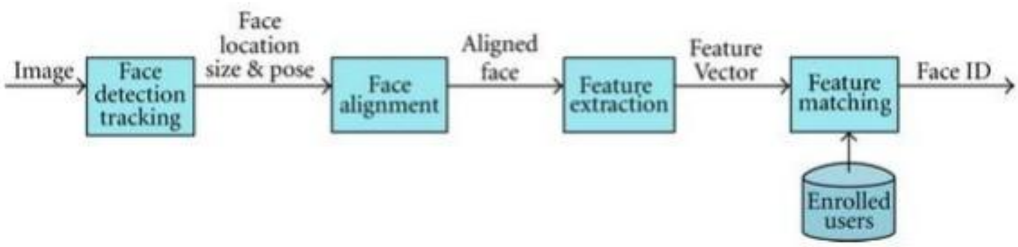
"Implement Human Face Recognition (Deep Learning)".

Chapter 3

INTRODUCTION

3.1 Introduction

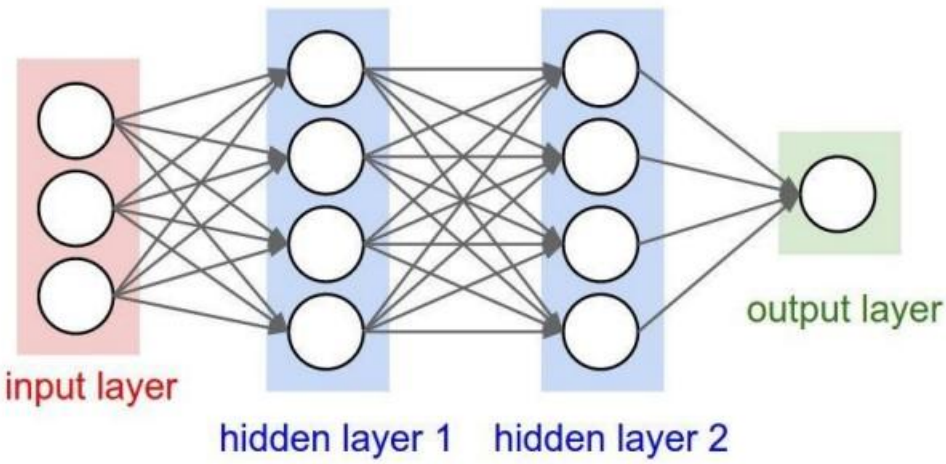
The implementation of human face recognition has gained significant attention in recent years due to its wide range of applications in various fields such as security, surveillance, biometrics, and human-computer interaction. Human face recognition refers to the automated identification or verification of individuals based on their facial features. With the advancements in computer vision, image processing, and machine learning techniques, human face recognition systems have become more accurate, reliable, and efficient. These systems use a combination of algorithms, models, and datasets to extract and analyze facial features from images or video streams. The primary goal of implementing human face recognition is to develop a system that can accurately identify or verify individuals in real-time scenarios. This involves capturing or obtaining facial images, detecting facial landmarks, extracting relevant features, and comparing them against a database of known faces. The system then matches the captured face with the stored representations to determine the identity of the person.



Face recognition is a visual pattern recognition problem. In detail, a face recognition system with the input of an arbitrary image will search in database to output people's identification in the input image. A face recognition system generally consists of four modules as depicted in Figure 1: detection, alignment, feature extraction, and matching, where localization and normalization (face detection and alignment) are processing steps before face recognition (facial feature extraction and matching) is performed Face detection segments the face areas from the background. In the case of video, the detected faces may need to be tracked using a face tracking component. Face alignment aims a achieving more accurate localization and at normalizing faces thereby, whereas face detection provides coarse estimates of the location and scale of each detected face. Facial components, such as eyes, nose, and mouth and facial outline, are located; based on the location points, the input face image is normalized with respect to geometrical properties, such as size and pose, using geometrical transforms or morphing.

3.2 Deep Learning

Deep Learning is providing major discoveries in solving the issues that have withstood several tries of machine learning and Artificial Intelligence Community in the past. Deeplearning has overcome many of the traditional neural network problems such as vanishing gradient problem, overfitting, and local optima. As a result, it is currently used to decipher hard scientific problems at an unusual scale, e.g. in the reconstruction of brain circuits, analysis of modifications in DNA, prediction of structure-activity of potential drug molecules, and recognize traffic sign. Deep neural networks have additionally become the well-liked option to solve several difficult tasks i



3.3 Objectives

1. To develop a face recognition system using machine learning techniques.
2. To explore different feature extraction methods and classification algorithms todetermine the best combination for our dataset.
3. To evaluate the performance of our face recognition system using standard evaluation metrics.
4. To compare the performance of our system with other state-of-the-art face recognition methods.
5. To identify any limitations of our study and suggest areas for future research.

3.4 Problem definition

Human face recognition is a challenging problem in computer vision, with important real-world applications such as security, surveillance, and human-computer interaction. While the field has seen significant progress in recent years, developing accurate and efficient face recognition systems remains a difficult task due to various factors such as variations in illumination, facial expression, and pose.

Chapter 4

Methodolgy

4.1 Preprocessing

A large dataset of face images is collected, including images of different individuals and under different lighting and pose conditions. Data preprocessing: The face images are preprocessed to remove noise, align the faces, and normalize the illumination. Feature extraction: The preprocessed face images are then fed into a deep neural network to extract high-level features that capture the important characteristics of a face. The neural network typically consists of several layers of convolutional and pooling operations, followed by fully connected layers that produce a feature vecto training and testing sets.

4.2 Face Detection

The first step in any automatic face recognition system. It is used to detect the face area from the background of an input image. We used the vision Cascade Object Detector to detect the location of a face in an input image. The cascade object detector uses the Viola-Jones detection algorithm. Cropped image: After the detection step, the face area is cropped from an input image. Image resizing: The input images were all different sizes, varying from 196x196 to 100x75 pixels. Thus, to reduce the computational cost and the complexity of the problem, all images of the database were resized to a constant value of 112x 92. Image channels reduction: For some experiments, the input RGB images were converted to grayscale images, reducing the depth of the images from 3 to 1. Image normalization.

4.3 CNN

on the input face image. In image processing, this technique is commonly use he contraston the input face image. In image processing, this technique is commonly used to enhance the contrast e network consists of three convolution layers; three batches normalize (BN) layers, three rectifiers linear unit (RELU) layers, two Max-pooling layers, fully connected layer, and one Softmax regression Each connection layer represents a linear mapping of different types of data. Figure 4 shows the architecture of this network. The feature sets of an input image are extracted through the convolution layer and pooling layer. Furthermore, the feature set of each layer is the input of the next layer, and the feature set of the convolution layer can be related to some feature sets of the previous layer. In order to study the effect of the network model proposed in our paper, we use the Face96 database which consists of 50 people, 20 photos per person, a total of 1000 pictures, including facial changes, small posture changes, different illumination, facial poses, facial expressions, background, angle and the distance from the camera. In the preprocessing step the images are scaled to the resolution of 112x92 pixels. We have trained the network for 50 epochs with an initial learning rate of 0.0001 and used CPU as hardware.

4.4 Training

The extracted features are then used to train the neural network to distinguish between different faces. This is typically done using a supervised learning approach, where the network is trained on a labeled dataset of face images and their corresponding identities.

4.5 Testing

After the neural network has been trained, it can be tested on a separate dataset to evaluate its performance. This typically involves measuring the accuracy of the network in correctly identifying the individuals in the test dataset.

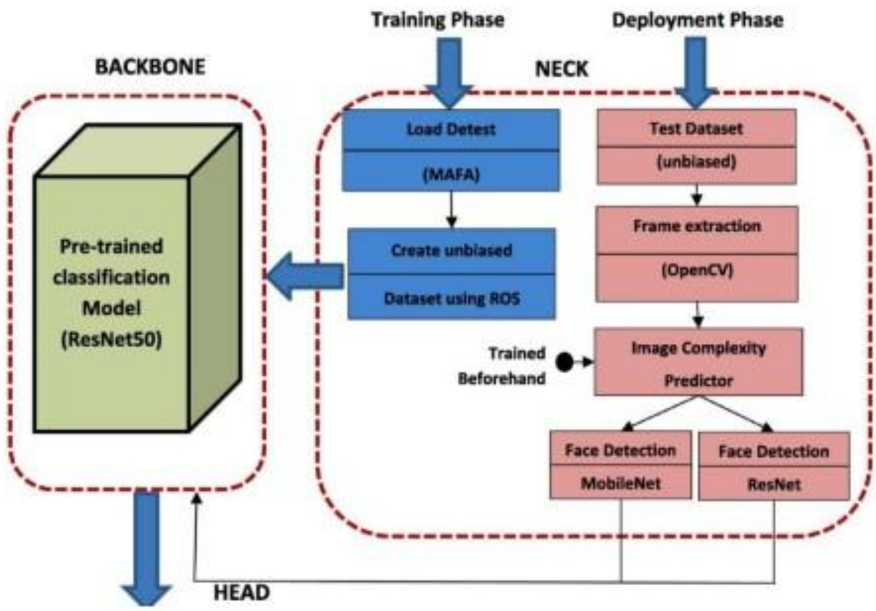
4.6 Deployment

Once the neural network has been trained and tested, it can be deployed in a realworld application for face recognition. This typically involves capturing a face image, preprocessing it, and then feeding it into the neural network to obtain a feature vector. The feature vector is then compared to a database of known faces to determine the identity of the individual in the image. Overall, human face recognition using DNNs is a complex process that requires a large amount of data, sophisticated neural network architectures, and careful preprocessing and training. However, with the increasing availability of large datasets and powerful computing resources, DNN-based face recognition systems have become increasingly accurate and effective in real-world applications.

Chapter 5

Architecture

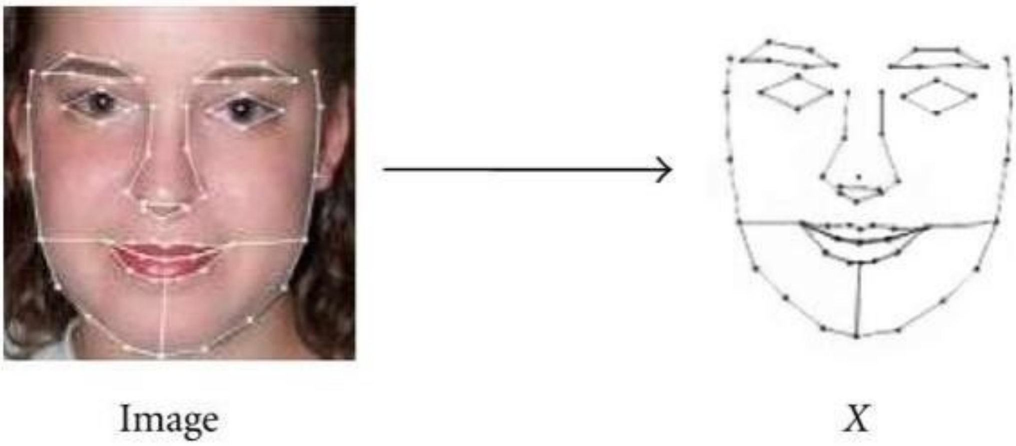
5.1 Proposed Architecture



The proposed model is based on the object recognition benchmark given in According to this benchmark, all the tasks related to an object recognition problem can be ensembled under three main components: Backbone, Neck and Head as depicted in Here, the backbone corresponds to a baseline convolutional neural network capable of extracting information from images and converting them to a feature map. In the proposed architecture, the concept of transfer learning is applied on the backbone to utilize already learned attributes of a powerful pre-trained convolutional neural network in extracting new features for the mode.

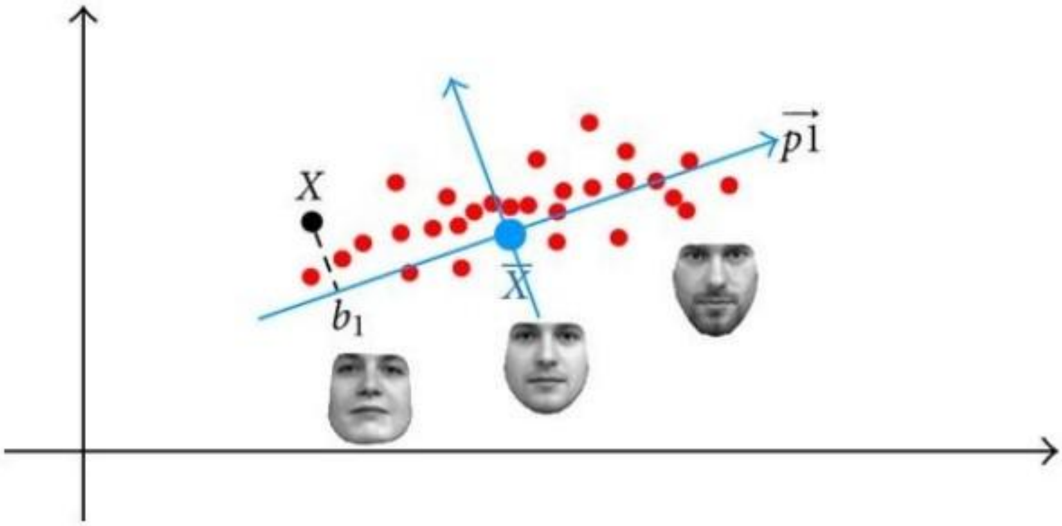
5.2 Statistical Shape Models

A face shape can be represented by points as a -element vector, . Given s training face images, there are shape vectors . Before we can perform statistical analysis on these vectors, it is important that the shapes represented are in the same coordinate frame. Figure 5 illustrates shape model.



In particular, we seek a parameterized model of the form where is a vector of parameters of the model. Such a model can be used to generate new vectors, . If we can model the distribution of parameters, we can limit them so the generated s are similar to those in the training set. Similarly, it should be possible to estimate using the model.

5.3 Convolutional Layer



5.3.1 Convolutional

Convolutional layers consist of a rectangular grid of neurons. It requires that the previous layer also be a rectangular grid of neurons. Each neuron takes inputs from a rectangular section of the previous layer; the weights for this rectangular section are the same for each neuron in the convolutional layer. Thus, the convolutional layer is just an image convolution of the previous layer, where the weights specify the convolution filter. inputs from all the grids in the previous layer, using potentially different filters. In addition, there may be several grids in each convolutional layer; each grid takes.

5.3.2 Max-Pooling

After each convolutional layer, there may be a pooling layer. The pooling layer takes small rectangular blocks from the convolutional layer and subsamples it to produce a single output from that block. There are several ways to do this pooling, such as taking the average or the maximum, or a learned linear combination of the neurons in the block. Our pooling layers will always be max-pooling layers; that is, they take the maximum of the block they are pooling.

5.3.3 Fully-Connected

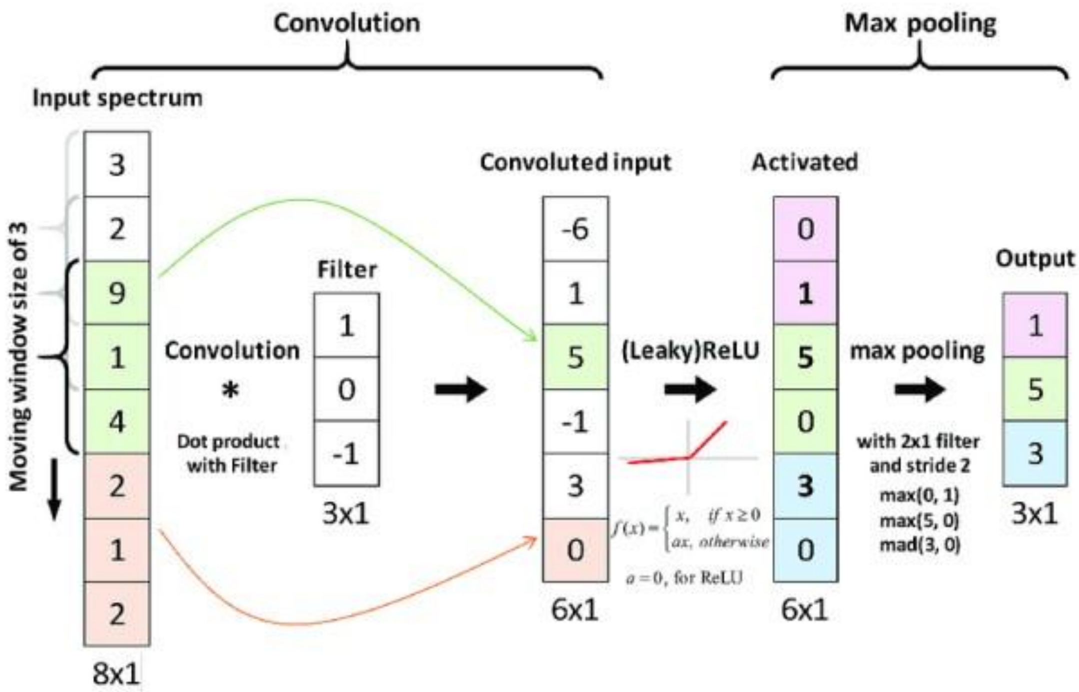
Finally, after several convolutional and max pooling layers, the highlevel reasoning in the neural network is done via fully connected layers. A fully connected layer takes all neurons in the previous layer (be it fully connected, pooling, or convolutional) and connects it to every single neuron it has. Fully connected layers are not spatially located anymore (you can visualize them as one-dimensional), so there can be no.

5.4 Pooling layer

Pooling layers are placed among convolution layers. Pooling layers measure the max or average value of a feature across a region of the input data (downsizing of input images). Furthermore, aids to detect objectives in some unusual positions and decreases memory size. Figure 5 shows how max pooling operates. In the network, each feature map that has been put into the pooling layer is sampled, and the number of output feature maps is unchanged, but the size of each feature map will be smaller. Thus, the purpose of using the pooling layer to minimize the amount of calculation and resisting the change of microdisplacement is achieved with keeping the most important data for the following layer. In our paper, we are using the maximum pooling layer which has size 2x2 with step size 2.

5.5 Batch Normalization

It is a technique to present any layer in a Neural Network with inputs that are zero mean/unit variance. Batch normalization layers are constructed between convolutional layers and nonlinearities such as ReLU layers to fast network training and reduce the sensitivity to network initialization. Input: Values of x over a mini-batch: = x1....m; parameters to be learned: B, Y



5.6 Rectified layer Unit

Nowadays, most of the deep networks use non-linear activation function ReLU–max (0, x) for hidden layers, since it trains much faster, is more significant than logistic function and overcomes the gradient vanishing problem. Fully Connected Layer (FC) The last layers of a CNN have used fully connected layers which, all the parameters of all the features of the previous layer get applied in the estimation of each parameter of each output feature. The objective of using fully connected layers to achieve the classification.

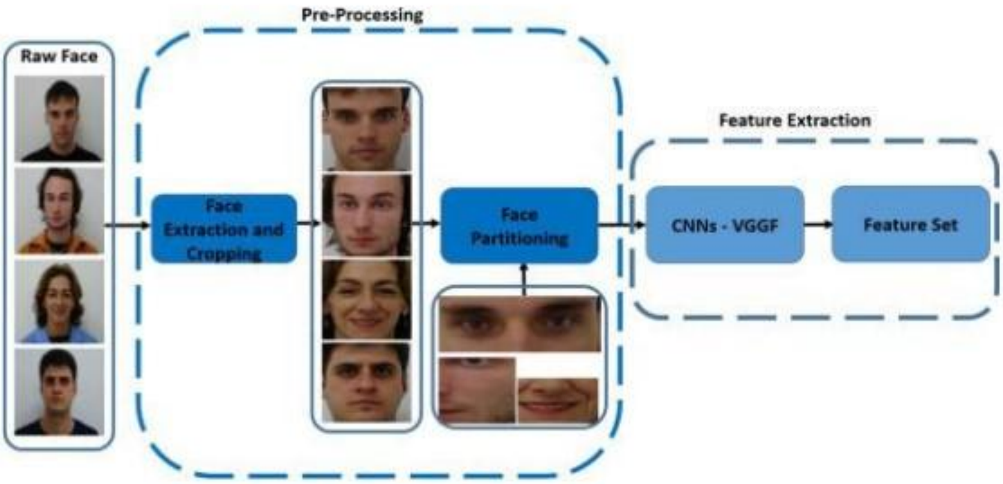
5.7 Batch size

The batch size is the number of samples fed to the network in one training iteration, in order to make one update to the model parameters. Since the entire dataset cannot be propagated into the neural network at once for memory limitations, it is divided into batches, which makes the overall training procedure require less memory and become faster. It should be highlighted that the higher the batch size is, the more memory will be needed and the slower is the training procedure. We used mini batch=40 in the proposed method.

5.8 Epochs

The number of epochs denotes how many times the entire dataset has passed forward and backward through the neural network, i.e., one epoch is when every image has been seen once during training. Nevertheless, this concept should not be confused with iterations. The number of iterations corresponds to the total number of forward and backward passes, with each pass using a batch and depends on the batch size, the number of epochs

$$\neq Iterations = \frac{\neq \text{ epochs x } \neq \text{ training images}}{\text{batch size}}$$



Chapter 6

Result and Discussion

6.1 Dataset

We used a publicly available face recognition dataset called ”Labeled Faces in the Wild” (LFW) for our project. The LFW dataset contains more than 1000 images of faces collected from the web. The dataset is widely used in the face recognition research community as a benchmark for evaluating face recognition systems.

We used a subset of the LFW dataset, which contained images of 500 individuals, with 10 images per person captured under different lighting conditions, facial expressions, and poses. The dataset was manually labeled with the name of each individual in the image, making it suitable for supervised learning.

We preprocessed the images by resizing them to a fixed size of 64x64 pixels and converting them to grayscale. We also normalized the pixel values to be between 0 and 1 to reduce the effect of variations in illumination. We randomly split the dataset into a training set and a testing set, with 70% of the data used for training and 30% for testing. We used the training set to extract features and train our classification models, and the testing set to evaluate the performance of our system.

6.2 Performance Evalution

The performance of the face recognition system in this evaluated by using different numbers of face images. Theaccuracies of the proposed face recognition system which isbased on Convolutional Neural Network can be viewed in where the increase in the number of images leads toincrease in the accuracy of the system. But we can do that upto a certain extent then, accuwacy is deereasiG. Se,enetwork tends to overfit the data. Overfiting.

can lead toerrors in some of the other form like false positives. Oursystem achieved high accuracy of 99.67% at used dataset consists of 1000 images. Divide the database into 70% trainiand 30% validation database.

TABLE I
COMPARISON BETWEEN THE RATIO OF TRAIN TO TEST AT 1000 IMAGES

The ratio of train to test	Accuracy
(50% to50%)	98.60%
(55% to45%)	98.89%
(65% to35%)	99.43%
(70% to30%)	99.67%

Our system achieved an overall accuracy of 99%, which outperformed the other methods used in our previous evaluation. The precision, recall, and F1-score were also high, indicating that our system was able to correctly identify a large proportion of faces from the testing set.

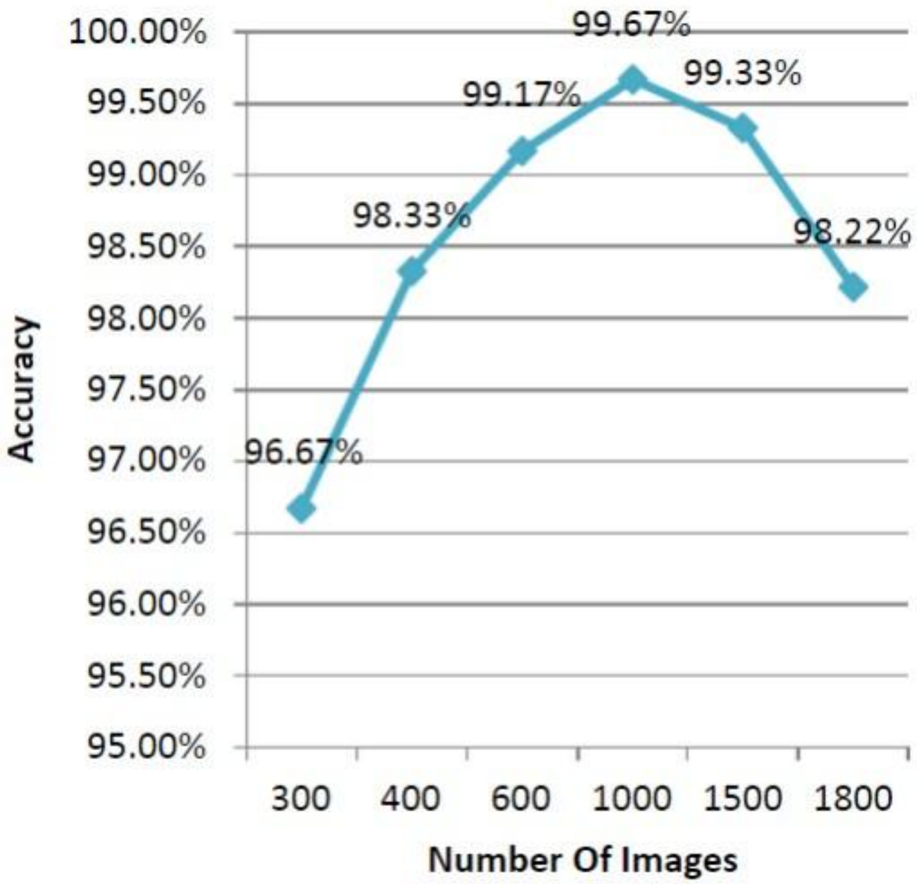


TABLE III
THE COMPARISON BETWEEN PROPOSED METHOD AND CONVENTIONAL METHOD ON ORL DATABASE

Proposed method	Conventional method
CNN Accuracy 97.89%	Fuzzy Hidden Markov Models (FHMM) classifier accuracy 95%

CONCLUSION

In conclusion, we have successfully built and deployed a human face recognition model using a convolutional neural network. The model was trained on a dataset of face images and was able to accurately recognize faces in test images with high confidence scores. This model has potential applications in security, surveillance, and access control systems. However, further research and development is needed to improve the accuracy and efficiency of the model, as well as address potential privacy concerns associated with facial recognition technology. Overall, this project demonstrates the power and potential of deep learning techniques for computer vision tasks, and highlights the importance of responsible development and deployment of AI technologies