CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Implementation of Neural Network for Effective Border Control

by

Farhat Yamin

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Computing
Department of Computer Science

2023

This thesis work is wholeheartedly dedicated to my beloved Mother Qamar Un Nisa, Husband Syed Ahmad Ali, Daughter Ayesha Khan and Zaina Khan who have been my source of inspiration and gave me strength when I thought of giving up, who continuously provided their moral support. Special thanks to my supervisor Dr."Masroor Ahmed" whose confidence enabled me to complete this milestone.

CERTIFICATE OF APPROVAL

Implementation of Neural Network for Effective Border Control

by

Farhat Yamin

(MCS191041)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Majid	UET, Taxila
(b)	Internal Examiner	Dr. Abdul Basit Siddiqui	CUST, Islamabad
(c)	Supervisor	Dr. Masroor Ahmed	CUST, Islamabad

Dr. Masroor Ahmed
Thesis Supervisor
April, 2023

Dr. Abdul Basit
Head
Dept. of Computer Science
April, 2023

Dr. M. Abdul Qadir
Dean
Faculty of Computing
April, 2023

Author's Declaration

I, **Farhat Yamin** hereby state that my MS thesis titled "**Implementation of Neural Network for Effective Border Control**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Farhat Yamin)

Registration No: MCS191041

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Implementation of Neural Network for Effective Border Control**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Farhat Yamin)

Registration No: MCS191041

Acknowledgement

In the Name of Allah, the Most Merciful, the Most Compassionate, Alhamdulillah all praises belongs to Almighty Allah, the Lord of the worlds and prayers and peace be upon Muhammad His Beloved Nabi and messenger. First and foremost, I must acknowledge my thanks to Allah, the Ever-Magnificent for His help and blessings. I am thankful to all who has significantly and insignificantly help to get my work done. I would like to take this opportunity to say warm thanks to all my loyal friends, who have been remained supportive along the way of completing my thesis. Last but not least, I would like to express my wholehearted thanks to my family for their support they provided me throughout my entire life and particularly through the process of pursuing the master degree. Because of their unconditional love and prayers, I have the chance to complete this thesis.

(Farhat Yamin)

Abstract

This paper includes the study about facial recognition technology and border control management processes. The facial recognition technology uses artificial intelligence based neural network which gets trained through appropriate classified data set. The proposed methodology explains that how the trained neural network can integrate in border control management processes to achieve the improved accuracy for face matching.

The proposed model is developed by defining the multiple layers of neural network which includes input layer, hidden layer, distance layer, dense layer, and output layer. The combination set of Max-pooling and Convolution ReLU is applied with channels and filters at multiple steps. Which gives binary classified results of matching and unmatching of image.

The matching data pools are defined in three distributions i.e Positive images, Anchor images and Negative images. The developed neural network is trained with the recognized data set for machine learning. The anchor distribution image is matched with Positive and Negative distributions and found 97-98% accurate matching.

Proposed Neural network can be considered useful to implement at Border control immigration process for real time or video processing for face verification.

Contents

List of Figures

List of Tables

Abbreviations

1:1	Refers to the relationship between two images A and B in which one element of image A may only be linked to one element of image B, and vice versa
1:N	Refer to the relationship between two images where there are multiple images to compare with single image
ANN	Artificial neural network
CNIC	Computerized National Identity Card
CNN	Convolution neural network
EPOCS	A greek word which means 'Fix point'. In english it is used to describe delay in date, hour, minutes or seconds
FRS	Face Recognition System
HOR	Histogram of oriented gradients
NADRA	National Database and Registration Authority
OCR	Optical character recognition
RFID	Radio frequency Identification

Chapter 1

Introduction

Facial recognition systems are become the need of the day to perform face authentication and highly demanding technological feature in intelligent applications. This is an advance version of biometric technology. Facial recognition technology can be used to perform efficient public safety, access control to important assets, suspect notification alerts, border surveillance, forensic investigations, crime detection, information security, and many others. Facial recognition implementation and operational challenges are increasing with its growing importance which includes subtle change in illumination, unpredictable poses, facial expression, aging and uncovered degree of face turning.

Face detection is first stage of face recognition process, in which the human face is identified from the whole image. The identification of edges of human face [1] are done through detection algorithm by identifying human face edges from image. These edges further measured the distance between face and head. Many researchers have proposed their algorithms to detect human faces with better degree of accuracy and speed.

Figure 1.1 shows the computer processing of face recognition system in which image is preprocessed for Face Detection. The algorithmic information about face measurements which include the size, color, gap between eyebrows, incline level of eyes is called face encodings. These encodings are the set of feature extractions of a face which are compared or identify the availability of the face intending to identify. Here is a general flow of Face Recognition technology: -

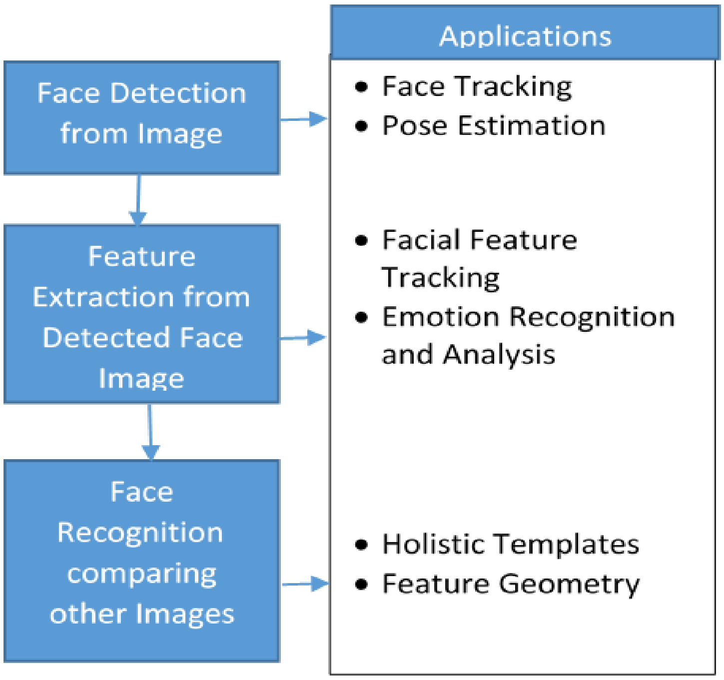


FIGURE 1.1: Face Recognition Contextual Process

The above figure shows that the face detection processing is applied to track a face and estimate the angle of the face in video feed. The extracted face features are further used to assess the emotion analysis and feature extraction. The extracted features are used to recognize the face by applying recognition techniques. The recognition of features is performed through holistic template or feature geometry approach. The holistic approach performs matching the features of face in portions which generate results with high accuracy. The feature geometry approach measures the facial points and curves between these data points to match face from one image to another image. These features include:

- Eyes
- Nose
- Lips

1.1 Face Detection

Face Detection is digital process to find the face in the processed image. That process includes are face edge finding, find the feature of a face and finding rest of features through computational analysis. One image can have more than one faces, so multiple approaches are used to detect all faces presence in image.

Following steps can be considered in face detection process:

1.1.1 Preprocessing

This step is performed to bring the image in appropriate format which can be help given model to find the face from image with high accuracy. The image preprocessing involves the resizing of the image, color, and illumination processing, setting up the orientation etc. The better preprocessing approach brings the neural model inference with better speed because it emphasis the features in an image in best possible way to detect and crop. And the giving specific pixels for matching bring accurate results in reduced time.

1.1.2 Classification

This is the process in which the unclassified features are added in their relevant categories. This is computer vision step to prepare the model for training. The good classification builds the neural model more efficient to bring results.

1.1.3 Localization

This is a process in which the model successfully identifies the edge of the object. The bounding box which is a boundary box to point out the presence of the object. The localization of the human face is improved to find the features of the face e.g. nose, eyes, chin etc.

1.2 Face Detection Approaches

1.2.1 Knowledge Based Approach

According to human visual system, face can be detected as a face if it has two eyes, one nose and lips and these features must have located in a relative distance and positions. The facial features can be extracted from an image as an input for further processes through knowledge-based methods.

This approach generally relied to identify face through image. But this is challenging that till what extent the rules can be applied. If these rules are too generic, then false positive come in output and if these rules are too details or strict then they may fail detecting the human face features from process image. Also, the pose of the human face is challenging to compute possible positive detection of features. Yang and Huang used a hierarchical knowledge-based method to detect faces [2]. Authors split the detection process into three levels. First level is the higher level generally describes that how the human face looks like. And lower level works on detailed features. As the figure 2 shows that how the human face points generally look like. Such representation availability is identified, and all available faces are taken to process one by one.

Figure 1.2 visually describes the process according to which image are dealt. Generally, the entire boxes (windows or grid). Each single box contains various levels of information.

FIGURE 1.2: General Share of Human Face

As figure 1.3 is supposed to be a detected face which is further checked through mosaic conversion of image (figure 1.4) to check the detailed analysis of features (Figure 1.5). The mosaic conversion process helps to identify facial features where every feature can be used for image matching.

Figure 1.3: Face detected from detection process [3]

Figure 1.4: Mosaic conversion for feature detection

Figure 1.5: An eye is detected from mosaic conversion as an object from multiple facial objects

Researchers has used the same approach to detect the face from complex background [4]. The complex background is processed using rule-based localization method [2]. Kanade [5] proposed an approach in which he used the vertical and horizontal assessment of the image to identify the position of the object. Then the left and right side of the head position was calculated to detect the boundaries of the face. The facial detection and recognition challenges includes the positioning of the face, and the variation is lifting up and bringing down the face angles.

Such issues have in in consideration of the researchers as well to find the algorithms and training the neural models to overcome these issues.

1.2.2 Feature Invariant Approaches Or Bottom-Up Feature Based Approach

This type of face detection methods is reverse approach to top to bottom approach discussed in section 1.2.1. The underline assumption is that the lighting condition and poses are not invariantly get changed/ appear in image because that would be challenging for this method to come up with accurate results. This approach detects the facial features and then infer the presence of a human face or multiple faces in the image. These facial features may include the eyes, eyebrow, mouse nose and hair line which is later used to detect the face edges/boundaries on the relative distance of facial feature points. Localization method [4] segments a face from a cluttered background for face identification. The approach is used by the edge mapping by [5] in which the group edges of multiple face image are removed to contour the available faces.

1.2.3 Template Matching Approach

In this approach a stander outline of a face is nailed down as sample or template. And processed image is matched where the related template outline/structure is present in the image. Actually, the existence of the human face is processed to detect on the basis of correlational value of template.

Multiple templates can be Used verity of templates of human features to train the model for human face [6]. Every template of the feature is pointed out with start and end point. These points are used to extracting the on the bases of change found in greatest gradient by matching with sub templates of each feature. The extracted contour template was matched with the first template and computation was made to calculate the relationship between both.

[7] Proposed the frontal view face which was able to present the localization approach using template which was used to apply in pre-defined environment under set pose and lighting conditions.

1.2.4 Appearance-based Approach

These methods are based on the statistical analysis and machine learning to identify the closest characteristics of a human face and tells the presence of face or no face. The statistical distribution model detects the presence of face.

The below figure shows the appearance-based face and no-face presence of multiple gaussian distributions. Gaussian distributions are the bell-shaped cure presence and the group of multiple gaussian distributions called the cluster. The first image is processed to detect the face samples by identifying the data points from grey scaled image as shows in 'Face sample Distribution'. The approximation of gaussian clusters presence is analyzed and used to generate the possible face features from the cluster. The features are shown as face centroids.

The figure 1.6 is processes to detect the face in which actually face is not present. The same above process is performed where face centroids shown the no face presence in the image.

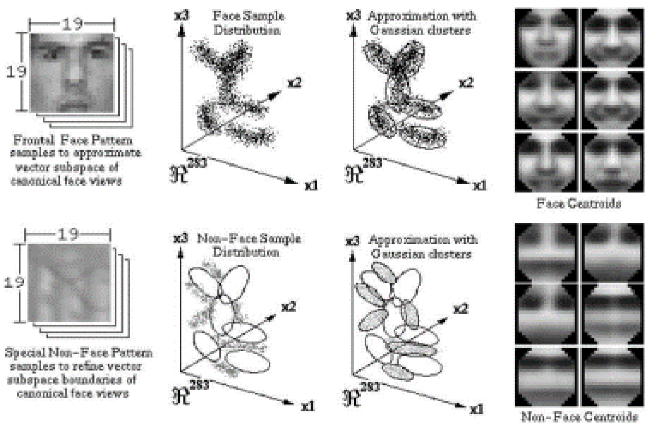


FIGURE 1.6: Face and nonface clusters used by Sung and Poggio. Their method estimates density functions for face and nonface patterns usinga set of Gaussians [1]

1.2.5 Eigen Faces Approach

Researchers provided an initial demonstration of eigen vectors being utilized in face recognition through a simple neural network that performed recognition of aligned and normalized face images. Kirby and Sirovich [8] suggested that technique for linearly encoding images of faces using a small number of basis images, which was first proposed by Pearson in 1901 and later by Hotelling in 1933.

The process begins by taking a collection of training images, represented as vectors of size mXn, and determining a set of basis vectors that span an optimal subspace. The goal is to minimize the mean square error between the projection of the training images onto this subspace and the original images. The basis vectors that accomplish this are known as Eigen pictures, as they are simply the eigenvectors of the covariance matrix computed from the vectorized face images in the training set.

Experiments using a set of 100 images show that a 91X50-pixel face image can be effectively encoded using only 50 Eigen pictures while still retaining a reasonable likeness, capturing 95% of the variance. This technique is a form of dimensionality reduction, as it allows for the compression of image data while retaining the essential features necessary for face recognition. It has since been used as a basis for developing more sophisticated face recognition algorithms that are capable of recognizing faces in real-world scenarios.

1.3 Neural Network

The neural networks are simulation of interconnected neurons inspired by natural brain. The capability is electronically established through feeding the possible scenarios. The deep learning mechanism is adapted to fed the model as the human learns and save the learning material which it is using as grow up. The biological signaling to interact with each neuron is built where input and output layers are connecting through hidden layers. These neurons are assigned the weights and when the input layers are sending the signals, the neurons are activated if the signaling brings up itself on threshold on output. In case there is output received not hitting the threshold then the data will not be moved to the next layer.

Neural networks comprise of two types i.e Artificial Neural Networks (ANN) and The Convolutional Neural Network (CNN) [9]. The Artificial Neural Networks are used to solve the complex problems and analytical reasoning. And Convolutional Neural Network are appropriate for computer vision problems.

CNNs are created on the deep learning problem solving approach. The architecture is used to perform data classification, Segmentation on the basis of features, face detections and recognitions and generalization. This CNN approach gives the high level of accuracy for computer vision problems and detections. And highly appropriate for facial recognition technology. Although the CNN architecture is based on the ANN architectural approach with more advanced and computer vision appropriate factors.

Artificial Neural Network is the artificial developed electronic model prepared through fed of possible scenarios and coordinates which make this model smart enough to match stored coordinates with ones needed to match.

The algorithm perform matching with speed. In other words, neural network could be considered a collection of programs that attempt to detect underlying links in a batch of data using a technique that resembles how the brain functions. In this context, neural networks are systems of neurotransmitters that might be biological or artificial in composition.

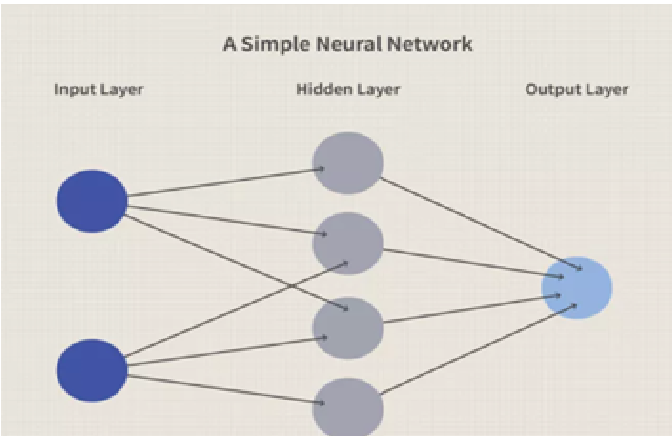


FIGURE 1.7: Artificial Neural Network Layers

There will always be data sets and task kinds that work well with pre-existing methodologies. The level of success of a neural network is ultimately determined by the well-prepared input data on the intended indication, not the algorithm.

The Convolutional Neural network becoming more popular by recognizing the recognition patterns received after detection process. The OCR, robot driving are the popular examples of CNN. Researchers has developed multiple algorithms for face recognition and detection processing in computer vision. These algorithms are the appropriate for detection only or recognition only. And many algorithms are equally offering the detection and recognition processes claiming the high accuracy levels. These algorithms have their owns good factors as well as challenges as well.

The CNN strength [10] include the difference between the filters applied and the layers used to compare the data gives the reliable results in computer vision. Where the input is divided into multiple chunks to match and process one by one as shown in figure below. The convolutional layer contains the filters which is smaller piece of a input processed image. These filters roll together to check the threshold. This threshold of the neuron weight enables the neuron to decide being activated or not [11].

The pooling is a process of down sample the output of neural network layer. The size of down sample and strid are used to process each sample. The pooling process calculates the value of each pixel where the minimum and maximum values can also use to speed up the process. The use of maximum value of image pixel is called the max pooling. The Max pooling layers assigns the value to each patch gathered from the feature maps which is used to take the sample from the feature map. Feature map is the organization of the neurons in a metrics for image processing [12].

The fully connected layer is the one where neurons are converted into the input vectors by taking the given weights. The threshold of each weight decides to activate or not.

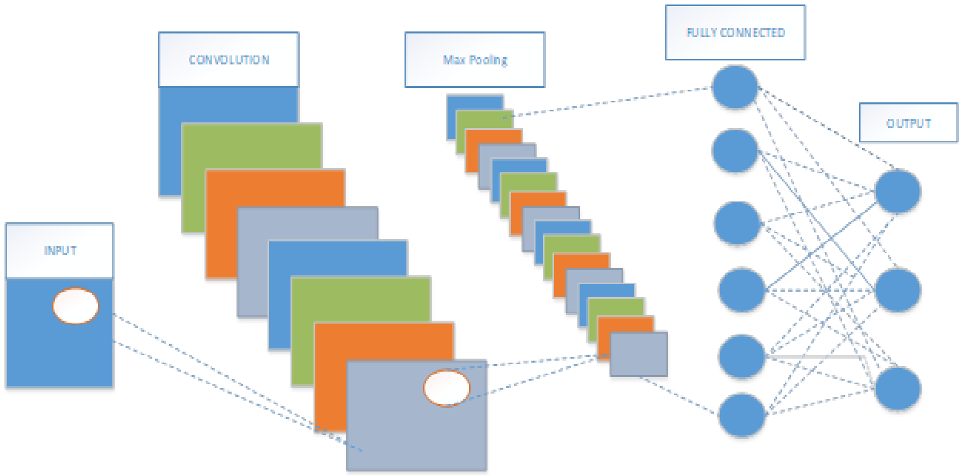


FIGURE 1.8: Convolutional Neural Network Architecture

The general proceedings of the data is shown in figure 1.8 in which the subnetworks used to apply filter along with the intensity value achieved after the training of model [8]. That unique value of the filter is derived the quality and type of the training set. That training set is shaped up as filter which is applied for the image detection. The first network of subnetwork applies the filters and the other subnetwork take the input from the first subnetwork to extract the feature value.

Researchers [13] has introduced the neural network consisting of 1024 units of inputs and 256 in hidden layers. There were hidden layers in which first layers were consisting of eight and the outputs were generating twins to bring the neuron activations.

Feraud and Bernier [11] has proposed the neural network in which five layers were used to perform mathematical computations to normalize the starting variables and calculate the correlations to create the feature vectors. There five layers were able to detect the frontal view of faces on the computations and turning that face image in 60 degree left and right. The facial detection and recognition process was strengthened using that model. Also, vigorous real time detection is researched by [14] P. V. a. M. J. Jones.

Goel [15] included the study of Lin et al in which the Probabilistic Decision-Based Neural Network (PDBNN) was proposed. This neural network was based on the feedforward neural network to use classification and pattern recognition. Later this approach was widely used in sensor setup and remote sensing image classification applications and tasking.

1.4 Facial Recognition

Facial recognition is the technological artificial capability developed to detect and recognize the facial feature of any human once it was registered before processing the facial coordinates for matching. This feature is primarily used for authentication the right person on any physical or logical data. Machine is providing the encodings to classify the types of measurements which are to calculate the distances between two faces [16].

In the face recognition problem, there are primarily two approaches that are commonly used: geometric or feature-based, and photometric or view-based. Over time, as interest in face recognition research has grown, various algorithms have been developed, and three of them have been extensively studied in the literature.

Face recognition is challenging due to variety of factors which includes:

1. In-appropriate lighting conditions
2. In-appropriate contrast level
3. Noise in image
4. Pose
5. Blur, etc.

These challenges are occurred low accuracy level in image detection and resultantly affect the facial recognition process [17]. The 3D model is used get the fair assessment of facial data points to overcome challenges mentioned above. Face Templates are also used to get the effective images recognition using bidirectional array of intensity values [10]. recognition is a subfield of computer vision and pattern recognition that involves identifying and verifying an individual based on their facial features. Neural networks are a popular machine learning technique that can be used to build facial recognition systems.

There are different types of neural networks used in facial recognition, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs). CNNs are widely used in facial recognition because they can detect features in images, including facial features. RNNs can be used to analyze sequences of facial images to recognize changes in expression or pose, while DBNs can be used to learn hierarchical representations of facial features.

Facial recognition neural networks are typically trained using large datasets of facial images, which are labeled with the identities of the individuals in the images. The neural network learns to identify key features in the images that distinguish one individual from another, and it can use this knowledge to identify new faces that it has not seen before.

There are several applications of facial recognition neural networks, including security and surveillance, personalization of user experiences, and healthcare. However, there are also concerns about the potential misuse of facial recognition technology, including privacy violations and biased algorithms.

Overall, facial recognition neural networks have shown promise in a range of applications, but continued research is needed to address ethical concerns and improve the accuracy and robustness of these systems.

Muge proposed the Eigen faces to detect the effective detection of human face from an image [18]. The basic idea behind eigenface technology is to represent faces as a linear combination of a set of basis images, called eigenfaces, which are derived from a set of training images. These eigenfaces are essentially the principal components of the distribution of facial images, capturing the most significant variation among the images. By projecting a new facial image onto the eigenfaces, one can represent it as a linear combination of the eigenfaces, which can be used to identify the person in the image.

The eigenface approach generally proceeds by taking into account various important and unavoidable steps for example it collect a set of training images of faces, usually in a controlled environment with consistent lighting, background, and pose. It preprocess the training images by normalizing them to a common size, and removing any irrelevant information such as background or non-facial features.

PCA is applied to the set of training images to obtain the eigenfaces. This involves calculating the covariance matrix of the training images, and then computing the eigenvectors and eigenvalues of the covariance matrix. The selection of the most significant eigenfaces, which capture most of the variation in the training images. Given a new facial image, project it onto the selected eigenfaces to obtain a set of weights that represent the image in terms of the eigenfaces. The weights are compared of the new image with the weights of the training images to find the closest match, and identify the person in the image based on the closest match.

1.4.1 Facial Recognition Approaches

As shown in Figure 1.1, the process of facial recognition includes the use of different face recognition algorithms. These algorithms are based on two general approaches:

1.4.1.1 The Geometrical Approach

This approach in face recognition relies on the spatial arrangement of facial characteristics or landmarks. As shown in below figure, it involves detecting and locating the primary geometric features of the face, like the eyes, nose, and mouth. Following this, various geometrical measurements such as the distances and angles between these features are used to recognize and classify faces.

Geometric implementation of the imaged object can be described with the help of Fig 1.9. By geometric description, we mean that geometric orientation of the pixels describes various contents.

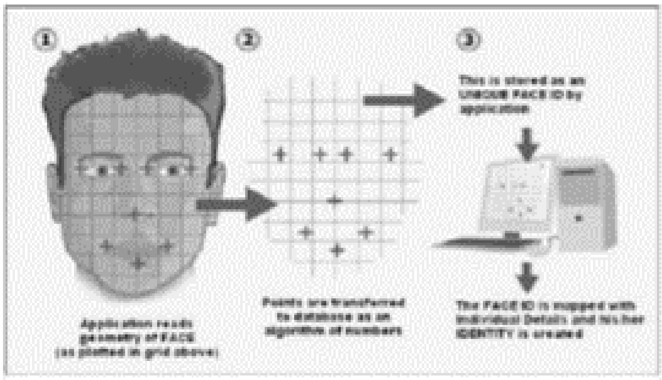


FIGURE 1.9: Geometric Facial Recognition [19]

1.4.1.2 Photometric Stereo

This method is utilized to reconstruct the shape of an object using multiple images captured under varying lighting conditions. The recovered object's shape is represented by a gradient map that comprises an array of surface normals, as described by Zhao and Chellappa in 2006 [20].

As shown in figure 1.10, the image present in left side is converted in the photometric image. The process is applied on the left image where the lighting conditions and direction are assessed on surface from right direction.

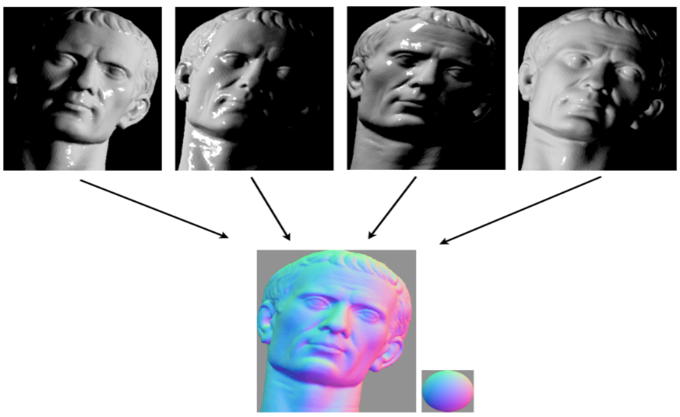


FIGURE 1.10: Photometric stereo [21]

1.5 Neural Network

The role of the neural network in terms of recognition, a given face processed according to subsequent details.

1.5.1 Image Frame

Generally, in surveillance system a video is made. This video can only be processed by decomposing it into frames. A frame is one of the numerous static pictures that make up the entire moving image. This refer to the single images which have been recorded on a strip of photographic film that gets longer in length while creating, from the beginning of modern filmmaking toward now, each image on such a strip looks rather like a framed picture when examined individually.

Below image shows the sample of image frames showing the car object with multiple pictures. These images are showing the presence of an object with slight distance to appear the movement from one place to another.

FIGURE 1.11: Video Stream Frames [22]

These frames are moving in sequence and one anther maintains the rythem in a way to give impression that the objects of frames are moving in actual. The speed is the frame rate in any video. In our research, the frame belongs to the single image taken to process for the purpose of face detection and recognition.

1.6 Challenges In Face Recognition

The face recognition is challenging job. And these challenges are expanding day by day as this becoming high in demand in variety of domains. The **pose variation** is the critical challenge in this technology which causes significant issues in face detection. The post variation happens due to change in face angle or head rotation. The facial recognition system is usually trained with the frontal face which is not enough to help the system recognizing processed face authenticity.

Image illumination is the lighting effect of the image which reduces the FRS efficiency. This challenges in sensitive in nature because the under illumination creates challenges to detect the object from the background whereas the over exposure or over lighting could be caused of poor face pattern detection. Multiple algorithm are used to resolve this challenge which may or may not be useful for the prevailing situation of the system.

The third challenges include the **variation of expression** of a face. As the nature is unlimited so this is not possible to capture every emotion and train the system for that.

Ageing is another important factor for facial recognition system where the human face remains unchanged in ten years but after the age of 65 the aging face doesn't get changed. So this is challenging for FRS to recognize face.

The occlusions is also create big challenge for FRS where the person wears the object e.h Cap, glasses or having bearded on face which creates issue recognizing the human face.

Facial recognition system has become the point to deep interest by Border Control organizations. They are incorporating the Facial Recognition system to monitor and control the human activities for better surveillance. Though the internal processes of Border Control are mix of Closed and open environment. The closed environment is the one where the traveler can be asked to give an image shot in set pose e.g the passport verification desk, information desk, the passenger security checking. Whereas the open environment includes the area where passenger is allowed to move and make the position naturally e.g. passing through, walking area, boarding and launching area.

The **close environment** contains illumination control issues, aging factor, change in style e.g beard, twins etc. The **open environment** contains close environment issues along with pose, blurriness, distance visibility etc.

1.7 Motivation

Border control process is a critical process to secure the country. Currently most of the border control immigration authorities are processing the passport verification process through manual verification or with human. Due to the heavy immigration flow and maintain the transparency, it is highly needed by border control immigration office to adopt facial recognition system in effective way.

The diversity of facial recognition technology can be used to increase the implementation of surveillance processes for better monitoring and regulating. This can be achieved once the technical capabilities and Border control processes are well studied. That study should be able to produce the methodology to implement Facial Recognition Technology to reduce Open and close environments challenges.

1.8 Problem Statement

To the best of my knowledge the research gap exists for developing an automated system that can reliable and effectively manage an enhanced level of the border.

1.9 Research Questions

Q1. How facial recognition technology approach can used enhancing border security and effectively integrated into the border control management process?

Q2. How modern computational technology can be used to introduce a reliable border control mechanism?

1.10 Proposed Solution

The proposed solution is based on the study of workflow of Border Control Process and diversity of Facial Recognition technological. This study highlights the appropriate Border control processes where Facial Recognition system can be implemented to bring more monitoring and control on authentication of traveler. The proposed Facial Recognition system accuracy is enhanced through the use of Neural Model which is trained with data set.

1.11 Objectives

Objective 1: Identify the border control process where facial recognition technology can be integrated.

Objective 2: Train neural network with complex face data set that bring higher accuracy in results through overcome the challenges in Pose variation, open and close environment and illumination?

1.12 Organization of Thesis

The remainder of the research is formulated as follows:

Chapter 1: The Introduction of the research topic and objectives with problem statement and technological adoption for successful presentation of outcome.

Chapter 2: Literature Review with Related work and Gap in research is discussed.

Chapter 3: This chapter discusses the proposed methodology to solve the questions described in Chapter 1.

Chapter 4: This chapter reveals the performance Evaluation of the results derived after implementation of proposed mode.

Chapter 5: This chapter assembles the research work and give conclusion with future work

Chapter 2

Literature Review and Background Study

There are more than 120 countries issuing biometric travel documents integrated with Radio frequency Identification (RFID) chip. This chip contains the traveler's information in it along with digital images and other biometric information about traveler [10] [23]. The facial recognition technology can use this information to authentic the traveler's information. The facial recognition technology is considered as technological initiative to secure the border control in more efficient way. The digital image is the primary object to perform face detection and recognition processing. There are multiple stages where the digital image is processed to detect and recognize the face. These stages require computational cost in terms of resources and time. Although the main goal to use the facial recognition, is to secure the border immigration but the time to spend for verification cannot be ignored while using the technological feature.

This study is performed to identify the facial recognition implementation at border control with maximum provisioning of accuracy to proceed with the facial recognition stages.

Digital image is a computer readable format of an image which is attained with the transformation of every pixel of image into 2Diemention array. JPEG format is the most used format of digital format which stands for Joint Photographic

Experts Group. Recent studies are often using this extension for preprocessing of the picture [8]. A high intensity image can consist on (0, 255) and [0, 65535].

This digital image is preprocessed to make able for face recognition by converting the pixels in 0 to 1. Which is a grey image to identify the facial outline.

The grey image is transformed to black and white format where the facial landmarks can be identified and process to detect the face.

Since this a long process so technological adoption is mandatory to adopt. A digital image array is the combination of picture elements which is called pixels and each pixel contain a set of information along which includes frequency of colors (channels R,G,B). The archived two dimensions is marked at X and Y scales for visual representation to show up an image.

2.1 Structure of Facial Recognition Technology

Facial recognition technology work is sequential in nature where the output of the previous process becomes the input of the next process. This sequence of processes can be divided into multiple steps.

Face detection technology is essential for applications like automatic lip reading, facial expression recognition, and face recognition [24, 25]. Face detection and recognition frameworks are nearly identical. The framework from Shang-Hung Lin's research is made up of two functional segments: a face image detector and a face recognizer.

The face detection process is a first step to execute the face recognition process. The human face is identified from the entire image. Many algorithms are used to identify the face presence in an image. The HAAR algorithm is the one which is used to identify the presence of face. This algorithm is widely used in many famous libraries like opencv and keras. This identification of edges of a human face is another approach widely used for face detection. Mr. David [1], for example, has proposed a method for detecting human face edges in colored images by varying the distance between the face and the background. The distance between the face and the head was also measured. This method was also capable of identifying the eyes in greyscale images. Many researchers have contributed algorithms for detecting human faces. That detection process is critical in the recognition process. Because the recognition process is unable to correctly recognize the face of incorrectly detected data points.

The face image detector searches the image for human faces and localizes them against the background. After a face has been detected or localized, the face data points are decoded and searches there decoded frames in possible positions. All possible positions face are chopped and scaling is recognized. Commonly the facial frame is aligned the fiducial markers (these markers are the included the corners of the eyes). Also the quality of illumination is get settled for better results.

The template extraction process [26] converts the facial images into the arrays of numbers which is considered as a unique identification of face for recognition process. The comparison is performed with other templates and score is calculated to get the confidence level. The score of confidence gets better means the probability of accuracy get high. The face detection approaches may be used as the cascading [13] or neural network learnings [11]. Here the cascading approach requires the a lot more computation but gives reliable results. The computation costs involves the time consumption and the resources utilization which is not applicable in [15] real time environment of face recognition technology.

Whereas the deep learning technology is the most reasonable for the scenario, this study cover i.e the border control. The deep learning approach is the multi-tasking learning [27] process which produces the results with high accuracy in less time. The deep learning process simultaneously estimates the ethnicity, gender and age.

In this model era, the deep learning approach has replaced the traditional approaches [28]. The deep learning or the neural network model has the ability to process the image in more efficient way than traditional face recognition approaches. The training data of better quality provides the better results producing the discriminatory feature extraction and calculates the differences between the face encodings. The unconstrained scenarios are the better controlled by the deep learning neural network or convolutional neural networks. The unconstrained environment is the one where the control on lighting, exposure, human posture and the speed of object movement is not in the control while capturing the frame. The border control is the combination of controlled and uncontrolled environment. Whereas the face recognition model which can work better in the uncontrolled environment has bright level of confidence of providing better workings in constrained environment.

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm that are commonly used for image processing and analysis. CNNs have the ability to automatically learn and extract features from the input data that are most relevant for the task at hand, such as face recognition. This is done through a series of convolutional and pooling layers that progressively extract more complex features from the input data.

In the case of face recognition, a CNN is trained on a large dataset of facial images, where each image is labeled with the identity of the person depicted in the image. During the training process, the CNN learns to identify the features that are most relevant for discriminating between different individuals, such as the shape of the eyes, nose, mouth, and other facial characteristics.

One of the key advantages of CNNs for face recognition is their ability to handle variability in facial appearance due to factors such as age, ethnicity, head direction, illumination, occlusions, and other factors. By training on a large and diverse dataset, the CNN is able to learn to recognize faces across a wide range of conditions, making it well-suited for real-world applications.

CNNs have become the standard for face recognition because they have been shown to achieve state-of-the-art performance on benchmark datasets, such as the Labeled Faces in the Wild (LFW) dataset. In addition, many commercial face recognition systems, such as those used for security and surveillance, are based on CNNs due to their high accuracy and robustness [29].

They use a huge number of annotated face snaps to train a learnable function made up of linear and nonlinear operators (typically millions). Recent methods have gradually improved performance on benchmark assessment datasets, mostly as a result of different architectural features and loss function selections [29][17]. For real-time deployment, the number of phases and parameters in the template extraction network is very important because it may significantly alter the network's processing needs.

Enrollment and verification/identification are the two separate processes in face recognition applications. [30] In order to store the generated templates in the gallery and link them to the appropriate subjects, the system is fed a collection of annotated face photographs during registration. The system is supplied with an unknown subject face at the verification or identification stage, and the generated template is compared to the gallery. In the verification scenario, the system is provided a face in addition to a supposed identity that needs to be verified. In the event of identification, the probing template is compared to each template in the gallery, and the system returns one or more candidates if the matching score goes beyond a predetermined limit [31].

The electronic arrangements of the passengers in and out is controlled through ABC (automated border control) gates to implement the facial recognition technology [20]. The ABC gates are adopted to speed up the facial recognition process without compromising any physical security threat. [32] Germany and USA are the best example of adopting the ABC gates to get better monitoring and control on their border Management Process.

The eligible passengers entering the United States who have the option to use Automated Passport control (APC) kiosks instead of filling out a paper Customs declaration form. The use of APC kiosks is intended to expedite the passport control process and provide a more efficient and user-friendly experience for travelers [33].

When eligible passengers arrive at the passport control area, they can proceed directly to the APC kiosks. At the kiosk, travelers will be prompted to scan their passport, take a photograph using the kiosk, and answer a series of CBP inspection-related questions verifying biographic and flight information. The questions are designed to assess the traveler's eligibility to enter the United States and to identify any potential security risks [34–37].

Once passengers have completed the series of questions, a receipt will be issued. The receipt contains a barcode that is used to identify the traveler's information and answers to the inspection-related questions. It is important that travelers keep the receipt as they will need it to finalize their inspection with a CBP Officer.

After completing the APC kiosk process, travelers will bring their passport and receipt to a CBP Officer to finalize their inspection for entry into the United States. The CBP Officer will review the information provided at the kiosk, verify the traveler's identity, and make a final determination about the traveler's eligibility to enter the United States [18].

One advantage of using the APC kiosks is that they allow people residing at the same address to be processed together. This can be particularly beneficial for families or groups traveling together as it can reduce the amount of time spent waiting in line.

Overall, the use of APC kiosks is a streamlined and efficient way for eligible travelers to enter the United States. It allows travelers to complete the necessary customs and immigration processes quickly and easily, while also providing CBP with the information they need to ensure the safety and security of the country.

The border control kiosk control and the suspect detection data has always been challenging for the immigration bodies. The implementation of the facial recognition demands the source data to match the real images. For that the suspect recognition has also been under consideration by all immigration authorities. Canadian immigrants are given the instructions to take the refuges facials biometric prior to the clearance [38].

As the Facial Recognition technology is able to define the facial ethnicity, so that could be utilized to verify the authenticity of provided credentials [38]. The evolving risks of border control will be doubled till 2035 which needs to improve the security features continually [39]. Improving the security risks doesn't only pointing to the data security but the physical security also threatening vastly. The Facial recognition has becoming the point of attention to use for mitigation of these potential risks. But the implementation and wise use of this technology is in demand to use it in an effective manner so that the processes of immigration doesn't get delayed nor the accuracy issues are raised [40][41]. The facial recognition technology is as much as the software efficiency demanding as the network needs to be efficient as well.

The potential power of multiple features of the Facial recognition technology is vast so this is another challenge for implementing entities to provide and manage the resources. The efficiency of the technology should be strategic to make possible the provisioning of the smooth services [42].

The research of border control and importance of facial recognition has always been considered for better border control but the required components and the flow of technology implementation seems needed to address more. The processes of border control are much more seeker of time and speed which still need to address and considered in researches.

The figure 2.1 shows the technology adoption survey [42] in which of travelers and border control processes. The booking part is the one used with high rate whereas the border control is still very low. Technology intensive countries are still need to adopt the technical featured effectively to take the advantage of Facial Recognition Technology. The common expectation from Faial Technology is not only the face identification and recognition but the more expectations are fixed while implementation which included ethnicity recognition, motion analysis and crowd detection along with notification and alerts etc.

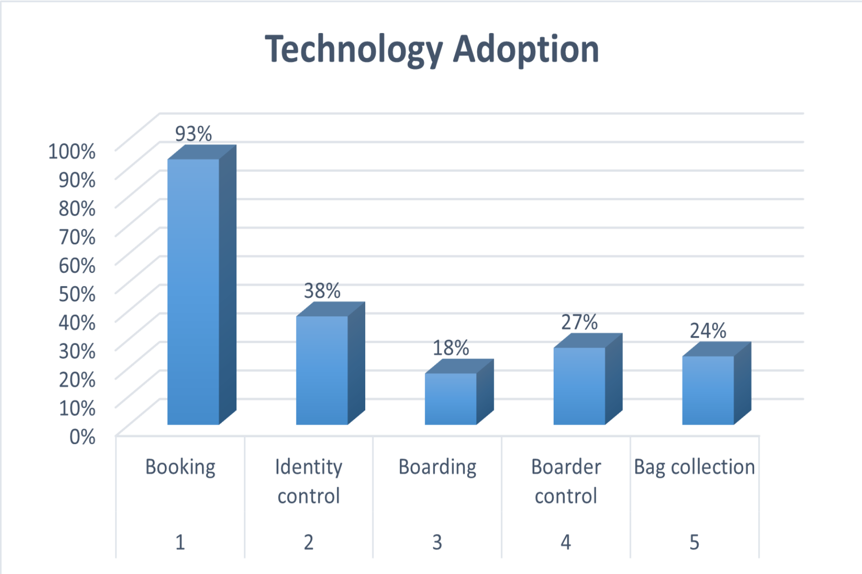


FIGURE 2.1: Technology Adoption for Travelers

The figure 2.1 describes that the travelers are adopting technological features to use plan and make their travel happen. The facial recognition adoption is at top in the process. The capture of traveler's face image can be used to feed and authenticate travelers face at initial stage. This will help to monitor and control the physical security from the beginning of travel process.

2.2 Face Detection Algorithms

Viola Jones is a framework for detecting faces proposed by Paul Viola and Michael Jones in 2001. It achieves high detection rate while rapidly processing images. It gives a rate of 15 frames per second. The algorithm is implemented in OpenCV. It has four stages

1. Haar feature selection
2. Creating an integral image
3. AdaBoost training
4. Cascading classifiers

2.2.1 Haar Feature Selection

The HAAR algorithm is a popular computer vision technique used for detecting objects in images. It was first introduced by Viola and Jones in their 2001 paper "Rapid Object Detection using a Boosted Cascade of Simple Features. "The HAAR algorithm works by detecting features in an image that are similar to patterns of dark and light regions known as "Haar features". These features are created by subtracting the sum of the pixel values in one rectangular region from the sum of the pixel values in another rectangular region.

To detect faces using the HAAR algorithm, a classifier is trained on a set of positive and negative examples of faces. The classifier is typically a boosted cascade of decision trees, where each tree consists of a set of HAAR features and a threshold value. During detection, the image is scanned using a sliding window approach, where the classifier is applied to each window at multiple scales. If a window is classified as containing a face, it is passed to a subsequent classifier for further verification. This process continues until a certain number of stages have been passed or the window is rejected.

The HAAR algorithm is widely used for face detection due to its high accuracy and speed. However, it can also have limitations such as being sensitive to variations in lighting conditions and orientation. Every human face shares some common properties, these properties can be matched using Haar like features. Few common features of human faces are:

- Eye region is darker than the nose bridge region.
- Upper cheek region is brighter than the eye region.
- Location of eyes, mouth, nose-bridge etc.
- Value = Oriented gradients of pixel intensities.

These three features are calculated by the algorithm and are then searched in the image. It starts with one pixel per feature and matches it with the entire window.

Value of each feature is calculated by subtracting the white region from the black region. Each feature gives one value. After that, two pixels are taken for each feature and is matched across whole window, it again gives one value.

2.2.2 Creating an Integral Image

Integral image-based face detection is a type of computer vision technology used to detect human faces in an image. The process involves several steps. First, the input image is converted to grayscale, and an integral image is calculated using a mathematical equation. Then, a sliding window is defined to scan the image, and the sum of pixel intensities within each window is calculated using the integral image. A pre-trained classifier is used to determine if each window contains a face, and the positions of all detected faces are recorded. Finally, bounding boxes are drawn around the detected faces, and the results are displayed. There are many ways to improve the performance of this technology, such as optimizing the size and stride of the sliding window and using more advanced classifiers.

2.2.3 AdaBoost Training

AdaBoost is a machine learning algorithm that is commonly used to train face detection models. The first step of this algorithm is to collect a set of positive face images and negative non-face images for training. Then, features are extracted from these images using techniques such as Haar-like features, which are rectangular features that measure differences in pixel intensities in different regions of the image.

A weak classifier is trained on a single feature to distinguish between faces and non-faces. This weak classifier is a simple classifier that performs only slightly better than random guessing. After training, the performance of the weak classifier on the training set is evaluated and the weights of the training samples are adjusted. Misclassified samples are given higher weights to force the classifier to focus on the more difficult samples.

Steps 3 and 4 are repeated with a new weak classifier and reweighting the training samples. The weak classifiers are then combined into a strong classifier using a weighted combination of their decisions. The weights of each classifier are determined by their classification accuracy.

The performance of the strong classifier is evaluated on a validation set and steps 3-6 are repeated until a satisfactory level of performance is achieved. Finally, the final strong classifier is used to detect faces in new images.

2.2.4 Cascading Classifiers

Cascading classifiers face detection algorithm, which was first introduced by Viola and Jones in 2001. The algorithm uses machine learning to divide the detection process into multiple stages, each of which uses binary classification to determine whether a particular region of an image contains a face or not. The classifiers are arranged in a cascade, with each stage processing a different feature of the image. The first stage quickly eliminates non-face regions using simple features, while subsequent stages use more complex features to refine the detection process and reduce false positives. The algorithm is efficient and widely used in applications such as video surveillance, human-computer interaction, and mobile robotics. However, it is sensitive to variations in lighting, pose, and occlusion.

2.3 Comparison of Facial Detection Algorithms

Here is a comparison table of face detection algorithms that utilize Haar features, integral images, AdaBoost training, and cascading classifiers:

Table 2.1: Comparison of Face Detection Algorithms

Algorithm	**Year**	**Speed**	**Accuracy**	**Hardware Requirements**
Viola-Jones	2001	Fast	Medium	Low
Skin color segmentation with Haar-like features	2003	Fast	Medium	Low

Algorithm	Year	Speed	Accuracy	Hardware Requirements
Table 2.1 continued from previous page				
Enhancing face detection using skin color segmentation and AdaBoost	2005	Medium	High	Low
Real-time face detection using modified skin color model and AdaBoost	2008	Fast	High	Low
Improving Haar-like features for real-time face detection	2010	Fast	High	Low
Face detection using Haar-like features and cascade of classifiers	2011	Fast	High	Low
Multi-scale Haar feature-based face detection using Cascade AdaBoost classifier	2013	Medium	High	Low

2.4 Face Recognition Algorithms

There are several face recognition algorithms that are commonly used in computer vision and machine learning applications. Here are some examples:

2.4.1 Eigenfaces

Eigenfaces is a face recognition algorithm that uses Principal Component Analysis (PCA) to represent faces as a linear combination of basis vectors or eigenfaces. The algorithm was first introduced by Sirovich and Kirby in 1987 and later extended by Turk and Pentland in 1991.

The Eigenfaces algorithm works by following these steps:

1. Collect a dataset of face images, each of which is preprocessed to have the same size and orientation.
2. Convert each face image into a vector by flattening it into a 1D array of pixel values.

3. Calculate the mean face by averaging all the face vectors in the dataset.

4. Calculate the covariance matrix of the face dataset by subtracting the mean face from each face vector, transposing the resulting matrix, and multiplying it by its transpose.

5. Compute the eigenvectors and eigenvalues of the covariance matrix. The eigenvectors represent the directions of maximal variation in the face dataset, and the eigenvalues represent the amount of variation in each direction.

6. Select a subset of the eigenvectors (i.e., the most significant ones) to form a set of basis vectors or eigenfaces. Typically, the number of eigenfaces is much smaller than the number of original pixels in the face images, which reduces the computational complexity of the algorithm.

7. Project each face vector onto the subspace spanned by the eigenfaces to obtain a set of feature vectors that represent each face in the dataset.

8. To recognize a new face, project it onto the same subspace and compare it to the feature vectors of the faces in the dataset using a distance metric such as Euclidean distance or Mahalanobis distance.

Eigenfaces algorithm has been widely used for face recognition, and it can achieve high recognition rates in constrained environments where lighting and pose variations are limited. However, it may struggle with more complex variations such as facial expressions or occlusion, and it may not perform as well in uncontrolled environments.

2.4.2 Fisherfaces

The Fisherfaces algorithm works by first computing the principal components of the face images using principal component analysis (PCA). This step reduces the dimensionality of the image data and extracts the most important variations in the images.

Next, the algorithm applies LDA to the PCA-reduced images to find the linear combination of features that maximizes the separation between different face classes while minimizing the variation within each class. The resulting discriminant features, called Fisherfaces, form a new basis for representing the face images.

During recognition, the algorithm projects a new face image onto the Fisherface space and compares it to the known face classes using a nearest-neighbor classifier or other distance-based methods. One of the main advantages of Fisherfaces is its ability to handle variations in lighting, pose, and expression, which are common challenges in face recognition. The algorithm can also be computationally efficient, especially when combined with PCA.

However, Fisherfaces has some limitations, such as its sensitivity to variations in illumination direction and its reliance on a small number of training samples per class. Researchers have developed various extensions and modifications to address these limitations, such as using more sophisticated feature extraction methods, incorporating local image descriptors, and using deep learning techniques. Overall, Fisherfaces remains a popular and influential algorithm in face recognition and has contributed to many important developments in the field.

2.4.3 Local Binary Patterns (LBP)

LBP is a feature-based algorithm that was first introduced in 1994 by Ojala et al. The algorithm works by converting the input image to grayscale and dividing it into small, non-overlapping blocks. For each pixel in each block, the algorithm compares its intensity value with its 8 neighboring pixels to the left, right, top, and bottom. Based on this comparison, the algorithm assigns a binary value of 1 or 0, resulting in an 8-bit binary code representing the local texture of the pixel. The algorithm then computes a histogram of the binary codes within each block and concatenates the histograms of all the blocks to form a feature vector that represents the texture of the entire image.

To detect faces in a new image, the LBP algorithm slides a window over the image and applies a binary classifier, such as a Support Vector Machine (SVM), to each block in the window. If most of the blocks in the window are classified as face regions, then the window is considered a face detection. The LBP algorithm has several advantages over other face detection algorithms, such as its simplicity, computational efficiency, and robustness to changes in lighting and facial expression. However, it may have a higher false positive rate compared to more sophisticated algorithms.

2.4.4 Deep Convolutional Neural Networks (CNNs)

Deep Convolutional Neural Networks (CNNs) have been successfully used in many applications, including face recognition. The key advantage of CNNs over traditional image recognition algorithms is their ability to learn hierarchical representations of images, which enables them to capture complex features and patterns.

In face recognition, CNNs are typically trained using large datasets of labeled face images, such as the Labeled Faces in the Wild (LFW) or the CelebA dataset. During training, the network learns to extract relevant features from the input images and map them to a high-dimensional feature space.

Once the CNN is trained, it can be used for face recognition by comparing the feature vectors of the input face image with those of the images in a database. This is typically done by computing the Euclidean distance or cosine similarity between the feature vectors.

There are several challenges associated with face recognition using CNNs, including variations in lighting, pose, and expression, as well as occlusions and variations in appearance due to aging. To address these challenges, researchers have developed various techniques, such as data augmentation, face alignment, and normalization.

Overall, CNNs have shown promising results in face recognition and are widely used in both academic research and industry applications.

2.4.5 Viola-Jones Algorithm

This algorithm is used for face detection but can be extended to perform face recognition by using the detected face regions as input to another algorithm. The algorithm uses a cascade of classifiers to quickly eliminate non-face regions and then applies a set of Haar-like features to detect faces.

2.5 Comparison of Facial Recognition Algorithms

Here is a comparison table of some popular face recognition algorithms:

Table 2.2: Comparison of Face Recognition Algorithms

Algorithm	**Year**	**Speed**	**Accuracy**	**Hardware Requirements**
Eigenfaces	1991	Fast	Medium	Low
Fisherfaces	1997	Fast	High	Low
Local Binary Patterns (LBP)	1996	Fast	Medium	Low
Viola-Jones	2001	Fast	Low-Medium	Low
Deep Convolutional Neural Networks (CNNs)	2012	Medium	High	Medium

2.6 Facial Recognition Technology on Border Control of Different Countries

Face recognition technology is being increasingly used in immigration control by various countries around the world. Here are some examples of face recognition models used on immigration control in different countries:

2.6.1 United States

The US Department of Homeland Security (DHS) has implemented a facial recognition system called the Traveler Verification Service (TVS) at various airports and ports of entry. TVS uses facial recognition technology to match the traveler's face against a photo on their passport or travel document.

2.6.2 United Kingdom

The UK Border Force uses facial recognition technology to verify the identity of travelers entering the country. The Automated Passport Control (APC) system, which is used at various UK airports, uses facial recognition technology to match the traveler's face with the photo on their passport.

2.6.3 Australia

The Australian Border Force uses facial recognition technology at various airports and seaports to verify the identity of travelers entering the country. The system uses biometric data such as facial images and fingerprints to match the traveler's identity with the data stored in the immigration database.

2.6.4 Singapore

The Immigration and Checkpoints Authority (ICA) of Singapore uses facial recognition technology at various checkpoints to verify the identity of travelers entering and leaving the country. The ICA's facial recognition system uses biometric data such as facial images and fingerprints to match the traveler's identity with the data stored in the immigration database.

2.6.5 China

China has implemented a facial recognition system called the "Smart Travel System" at various ports of entry to verify the identity of travelers entering the country. The system uses biometric data such as facial images and fingerprints to match the traveler's identity with the data stored in the immigration database.

2.7 Benefits of Facial Recognition Technology on Border Control

Facial recognition systems have the potential to provide a range of benefits when used in immigration control. Some of these benefits include:

Increased accuracy: Facial recognition systems can be highly accurate in identifying individuals, which can help to reduce errors and prevent cases of mistaken identity.

Enhanced security: By using facial recognition technology, immigration authorities can better identify and track individuals who may pose a security risk.

Improved efficiency: Facial recognition systems can automate many of the processes involved in identifying and verifying individuals, which can help to speed up immigration procedures and reduce waiting times.

Reduced fraud: Facial recognition technology can help to prevent identity fraud by verifying an individual's identity using biometric data that is difficult to fake.

Better data management: Facial recognition systems can help immigration authorities to manage large amounts of data more effectively and accurately, which can improve decision-making processes and reduce the risk of errors.

Overall, facial recognition systems have the potential to improve the accuracy, efficiency, and security of immigration control processes. However, it is important to balance these benefits against potential privacy concerns and ensure that the technology is used in a responsible and ethical manner.

2.8 Study on Border Control Workflow and Need of Facial Recognition System

Border control using facial recognition technology is a system that uses computer vision algorithms and machine learning techniques to identify individuals at border crossings based on their facial features. This technology has been gaining traction in recent years, as it offers a more efficient and secure alternative to traditional border control methods, such as manual passport checks and physical inspections [43].

The process of border control using facial recognition technology typically involves several steps. First, an individual approaches the border checkpoint, where their face is captured by a camera. The captured image is then compared against a database of known faces using facial recognition algorithms [44]. The database may include images of individuals who have previously crossed the border, as well as images of individuals who are on a watchlist for security or law enforcement purposes [43].

Border control process is a critical process to secure the country. Currently most of the border control immigration authorities are processing the passport verification process through manual verification or with human [18, 34]. Due to the heavy immigration flow and maintain the transparency, it is highly needed by border control Immigration office to adopt facial recognition system.

Our proposed system is highly viable for the border control to use as independent system or as an integrated API.

2.8.1 Border Control Management Process Context

- 1:1 Verification is used to authenticate foreigner's passport and visa information and grant access to them accordingly

- 1:N identification is used to create a watch list database who are resident but travelled abroad illegally by air, road and sea. Moreover, Interpol data (foreigner) is also integrated.

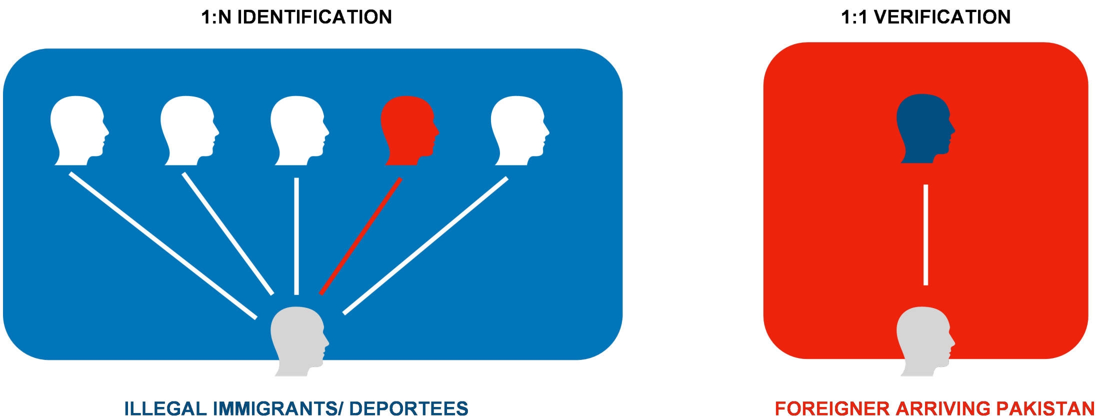


Figure 2.2: Identification Types at border

2.8.2 Traveler Journey

Generally, Passport is scanned on immigration officer desk to authenticate the traveler's credentials. The photograph of traveler is verified through naked eye and the current picture is saved to keep the record. The traveler is allowed for further processing upon successful verification.

The below figure 2.2 explains the four major steps along traveler journey. The traveler's passport is scanned to authenticate the document. The traveler's photo is captured in real time and that image is matched with the photo available on traveler's passport. The matching process is performed using facial recognition system. The authentication is performed and access to undergo next process is granted if authentication is successfully passed.

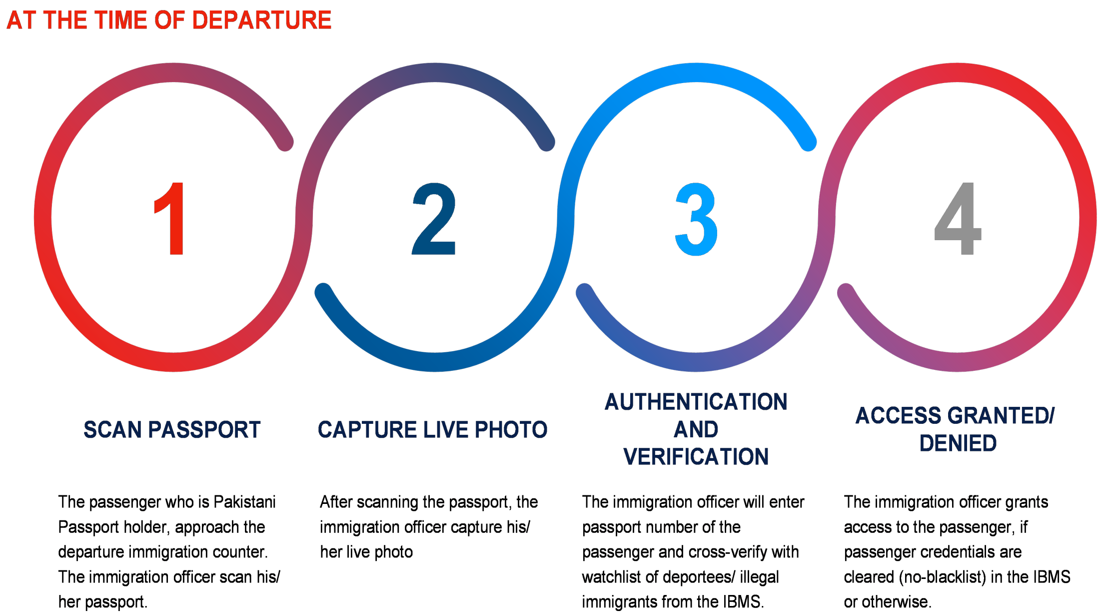


Figure 2.3: Traveler Journey

2.9 How Proposed Solution Work For Border Control Management System?

The proposed solution is able to integrate with existing system where the neural model is predict the authenticity of the traveler. Our proposed system is also capable to work as an independent system to capture the current image and verify to compare with registered face.

As shown in figure2.3, the live image is captured and the assessment is performed on traveler's live image and the information available with immigration system. If both information are not matched the system is able to generate alerts and notifications. The live monitoring is also performed on open environment where all travelers are monitored to identify suspect at any corner of boarder premise.

FIGURE 2.4: Proposed Solution Phases

The figure 2.4 explains the process flow point where the proposed system is proposed to use as an API. The authentication and verification process is the identified process area where the authentication and verification is performed of traveler. This process area is appropriate process to integrate the face recognition proposed model for the following reasons:

Increased Efficiency: Facial recognition APIs can automate the authentication and verification process, reducing the need for manual identity checks. This can improve the speed and efficiency of the process, reducing wait times for passengers and improving the overall airport experience.

Enhanced Security: By using a facial recognition API for authentication and verification, airports can ensure that only authorized individuals are granted access to restricted areas. This can help to enhance security measures and reduce the risk of security breaches.

Accurate Identification: Facial recognition technology has improved significantly in recent years and can now accurately identify individuals with a high degree of accuracy. This can help to reduce errors and prevent cases of mistaken identity.

Scalability: Facial recognition APIs can be integrated into existing systems and can be easily scaled up or down as needed. This makes it a flexible solution for airports that may experience fluctuations in passenger traffic and require a scalable solution.

Overall, integrating facial recognition APIs for authentication and verification processes can help to enhance the security, efficiency, and accuracy of airport operations, improving the overall traveler experience.

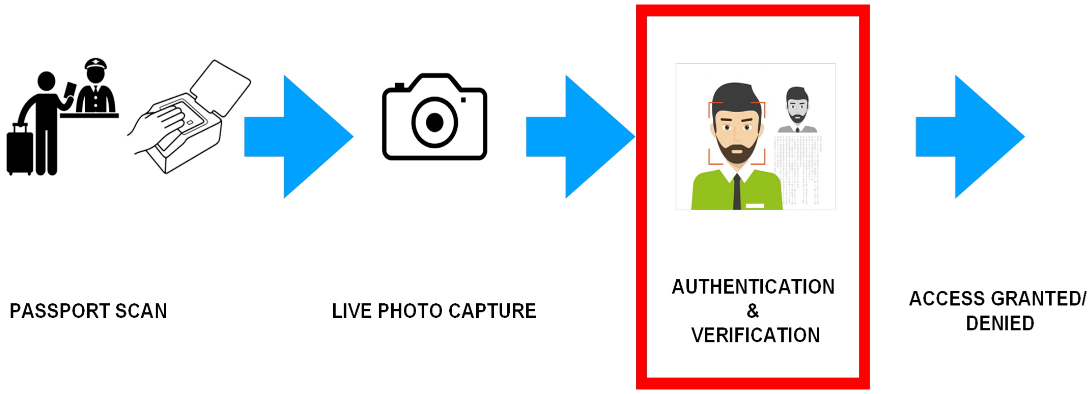


FIGURE 2.5: Integration Point of Proposed Solution in User Journey Process

2.10 Technical Model To Implement Border Control Using Neural Model

The immigration system is integrated with multiple types of analytical services where the whole data is collected and processed for certain decision making, alert management, Notification serving. The proposed system is suitable to integrate at that point of system where the processing of profile and verification is processed. The figure 16 is depicting the entire technical process where the proposed system fits in. The technical workflow of a proposed model to perform with face recognition functionality at border control typically involves the following steps:

- **Camera Placement:** The cameras are strategically placed in areas assigning the IPs where they can capture images of traveler's faces. These cameras may be fixed or pan-tilt-zoom cameras.

- **Image Capture:** The cameras capture images of traveler's faces and transmit them to the face recognition software for analysis. The images may be transmitted wirelessly or through a wired network.

- **Face Detection:** The face recognition feature analyzes the images to detect faces. As described in proposed methodology, it uses HAAR algorithm to detect facial features such as the eyes, nose, and mouth.

- **Face Recognition:** Once a face is detected, the CNN model compares it to a database of known faces to determine if there is a match. The database may contain images of individuals who are authorized to access the area being monitored or images of individuals who are on a watchlist.

- **Alert Generation:** If a match is found, the system can integration with alrm system to generate alert. The alert may trigger an action such as locking a door, sounding an alarm, or sending a notification to security personnel.

- **Database Management:** The face recognition system also manages the database of known faces. It may add or remove faces from the database based on changes in access authorization or watchlists.

- **System Maintenance:** Regular maintenance of the system is required to ensure that it continues to operate correctly. This may involve cleaning the cameras, data backup, improvement in ML according to evolving need or replacing hardware components

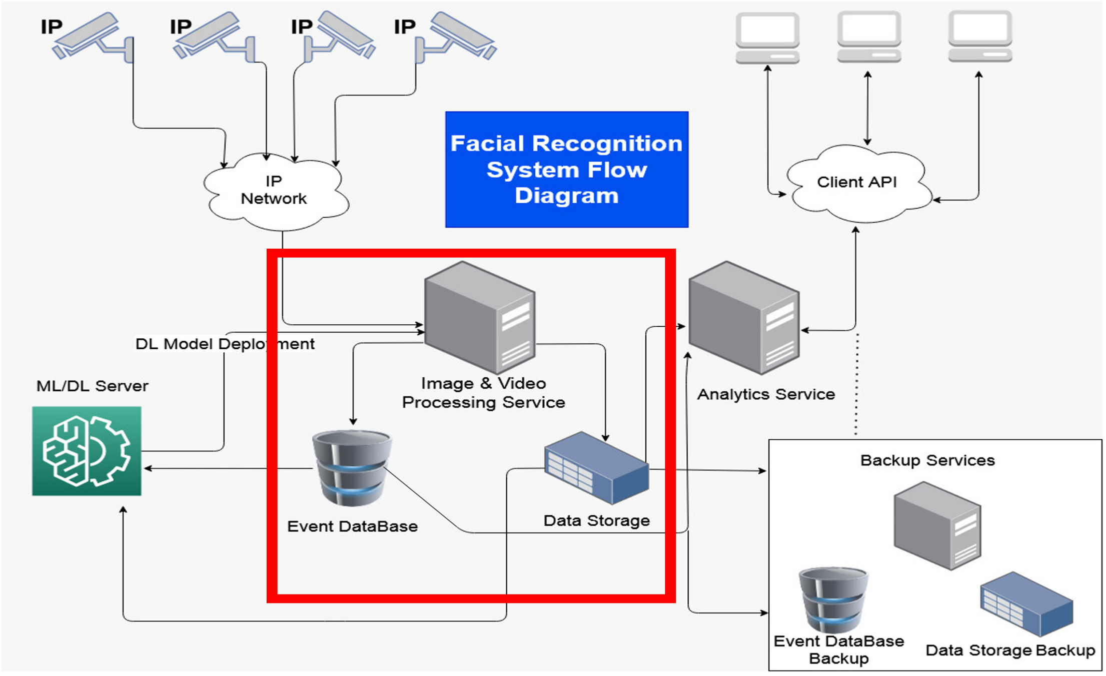


FIGURE 2.6: Technical Model of border control System to Integrate Proposed Solution

Chapter 3

Methodology

This chapter describe the methodology based on supervised learning approach to achieve the accuracy following less number of steps. The neural model is developed with six layers to process the images which is trained with the LFW data. The methodology is explained below.

3.1 Methodology

The proposed face recognition system is a type of biometric technology that is designed to identify individuals based on their facial features. As described in figure 3.1, this system is made up of two main components: the face detection module and the face recognition module.

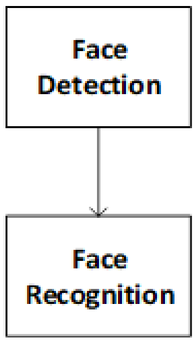


FIGURE 3.1: Proposed system Modules

The face detection module is responsible for locating and isolating the face from an image. This module uses various algorithms to detect the presence of a face in an image. Some common techniques used in face detection are discussed in literature review. The Haar cascades, Viola-Jones algorithm is selected to perform face detection. The predefined HAAR uses additional techniques to locate key facial features such as the eyes, nose, cheeks, chin and lips. These features are then used to align the face and normalize the image for the face recognition module.

The face recognition module is responsible for identifying the individual based on the facial features that have been detected by the face detection module. This module uses a combination of feature extraction and pattern recognition techniques to create a unique representation of the face. The Convolutional Neural Networks (CNNs) is selected to implement the facial recognition process.

Together, the face detection and face recognition modules form the technical capabilities required for a face recognition system to perform its expected processes. The system can be used for a variety of applications, such as security and surveillance, access control, and identity verification. The accuracy and reliability of the system depend on the quality of the images or video feed, the algorithms used in the face detection and recognition modules, and the size and quality of the database used for comparison.

3.1.1 Face Detection Module

The HAAR based approach is used to detect face in the image. The predefined detection module is used in our proposed model to perform face detection process. The implementation of HAAR approach is performed using OpenCv.

The OpenCV uses a specific type of Haar filter for face detection, which is called the Haar-like feature classifier. The Haar-like feature classifier is a type of cascade classifier that uses Haar-like features for object detection. Haar-like features are composed of rectangular regions with specific brightness patterns. These rectangular regions are placed at different locations and scales in an image to detect local features such as edges and lines.

The Haar-like feature classifier uses a trained model that consists of a set of Haar-like features and weights, which are used to determine whether a particular region of an image contains a face or not. As shown in figure 3.2, the classifier scans the image with a sliding window, and evaluates the presence or absence of each Haar-like feature at each location and scale. If a sufficient number of features are present in a given region, the classifier classifies the region as containing a face.

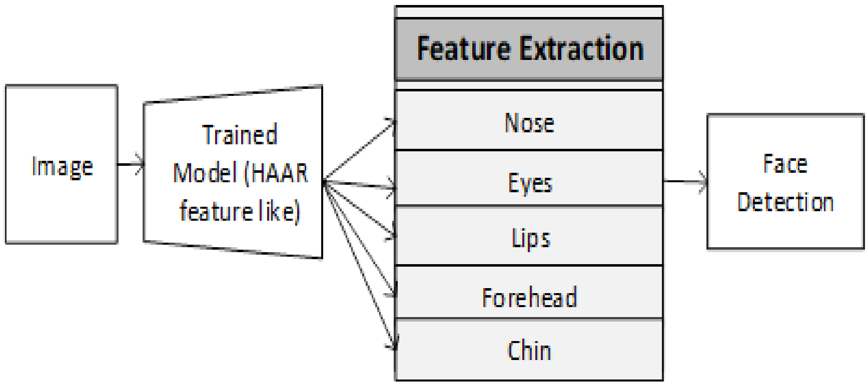


FIGURE 3.2: Face detection flow in HAAR like classifier

OpenCV provides a pre-trained Haar cascade classifier for face detection, which is trained using the Viola-Jones method. This cascade is applied in proposed model on large dataset of positive and negative examples to detect faces in a wide range of conditions, including different poses, facial expressions, and lighting conditions.

The algorithm detects the following features to identify the presence of a face:

Eyes: The Haar cascades algorithm identifies the eyes as two dark regions that are separated by a bright region (the nose). The algorithm looks for these features in the input image to detect the presence of a face.

Nose: The nose is also an important feature that the algorithm looks for to identify a face. It looks for a bright region between two dark regions (the eyes).

Mouth: The Haar cascades algorithm also looks for the mouth, which is identified as a curved, dark region below the nose.

Cheeks: The algorithm also looks for the cheeks, which are identified as rounded, bright regions on either side of the face.

Forehead: The algorithm also detects the forehead, which is identified as a flat, bright region above the eyes.

Chin: The chin is identified as a curved, dark region below the mouth.

The Haar cascades algorithm uses a combination of these facial features to identify a face in an image or video frame. The algorithm applies the features at different scales and sizes to detect faces of different sizes and orientations.

3.1.2 Face Recognition Module

The methodology for implementing face recognition functionality using involves several steps. The face recognition module is developed using CNN approach and trained using Keras and TensorFlow on the LFW dataset. The performance of the trained model is tested on a separate dataset of images that includes variety of faces. The matching of images is also performed to check the accuracy. The model performance is improved while testing by applying augmentation process.

The **augmentation** process is performed to enhance the data diversity for training. Images are flipped or rotated to create new examples that have different orientations and viewpoints. The scaling and cropping is performed to different sizes and aspect ratios, which help the model learn to detect objects at different scales and sizes. The brightness and contrast of images are also adjusted to simulate different lighting conditions.

Once the classifier and the recognition model are sufficiently accurate, the test system is developed using react at front end and embed this API at backend to test. The OpenCV library is used to capture video from a camera or load video files and apply the classifier and recognition model to each frame to detect faces and

FIGURE 3.3: Data Augmentation Samples [45]

recognize individuals in real-time. The facial recognition flow is described in figure 3.3 is described. The data is collected and split into the positive, negative data groups. The images are preprocessed and data augmentation is performed to get data ready for test and training. The image size is taken and the convolution layers is developed applying 64 filters. The maxpooling is applied on feature map and six layers are convolute as shown in figure. The fully connected layer performs the sigmoid function as non-liners activation function to calculate the feature vector and generate the output.

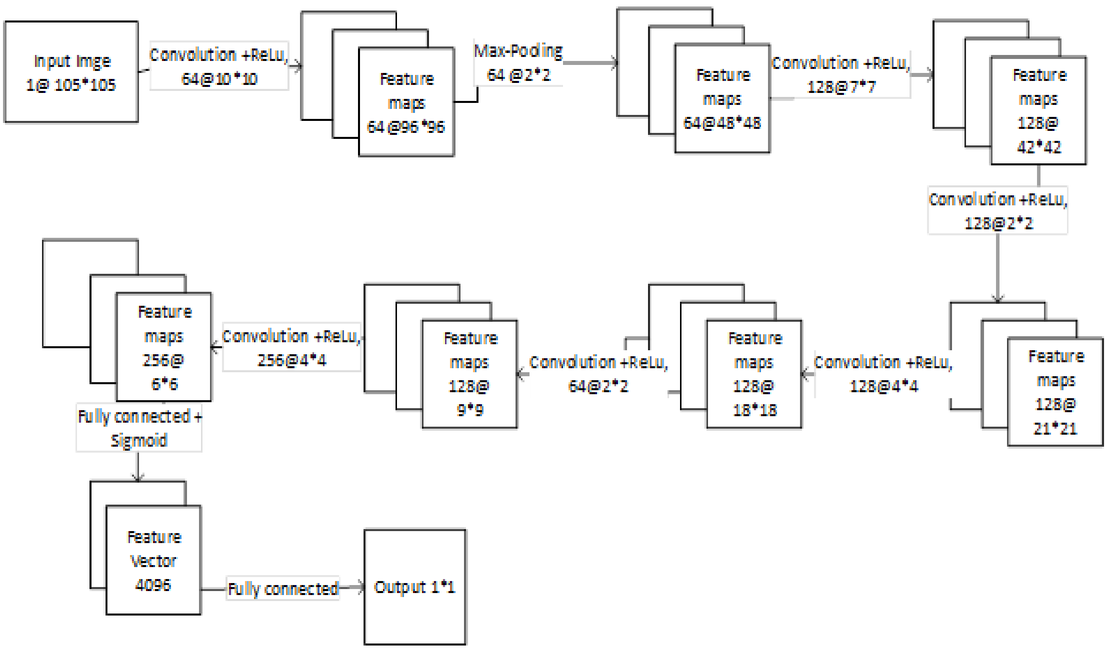


FIGURE 3.4: CNN layers of proposed model

3.2 Training

The model is trained to improve the learning of the model. The training process for a proposed neural network typically involves the following steps:

Data Preparation: The training dataset is prepared by pairing the input samples that are similar and dissimilar. For example, if the task is to identify similar images, then pairs of images from the same class are considered similar, while pairs of images from different classes are considered dissimilar.

Network Architecture: The proposed neural network architecture is designed and implemented. The network typically consists of two identical sub-networks that share the same weights. Each sub-network is responsible for processing one of the input samples and producing a feature representation of the input.

Loss Function: A triplet loss function is selected to train the network which helps the feature representations of similar samples to be close together and those of dissimilar samples to be far apart.

Training: The proposed network is trained using the 70% of training dataset and the selected loss function. The training process involves backpropagation of the error gradients through both sub-networks, and updates to the shared weights.

Validation and Testing: After training, the network is validated and tested on a separate dataset to evaluate its performance. The performance of the network is typically measured using metrics such as accuracy, precision, and recall.

The figure 3.4 shows the combined workflow of the proposed neural network face detection and recognition modules. The data set is collected and stored in data repository. The face detection module processes the image using HAAR algorithm where preprocessing is performed to resolve the brightness, color and size. The classifier detects the face (if available). Face recognition takes the detected face to match with the images available in data store. The face recognition is developed in CNN which is trained with 70% of data set where as 15% data is used to test and 15% of data set is used to validate the output. Once the outcome is received as expected. The results are displayed and process is completed.

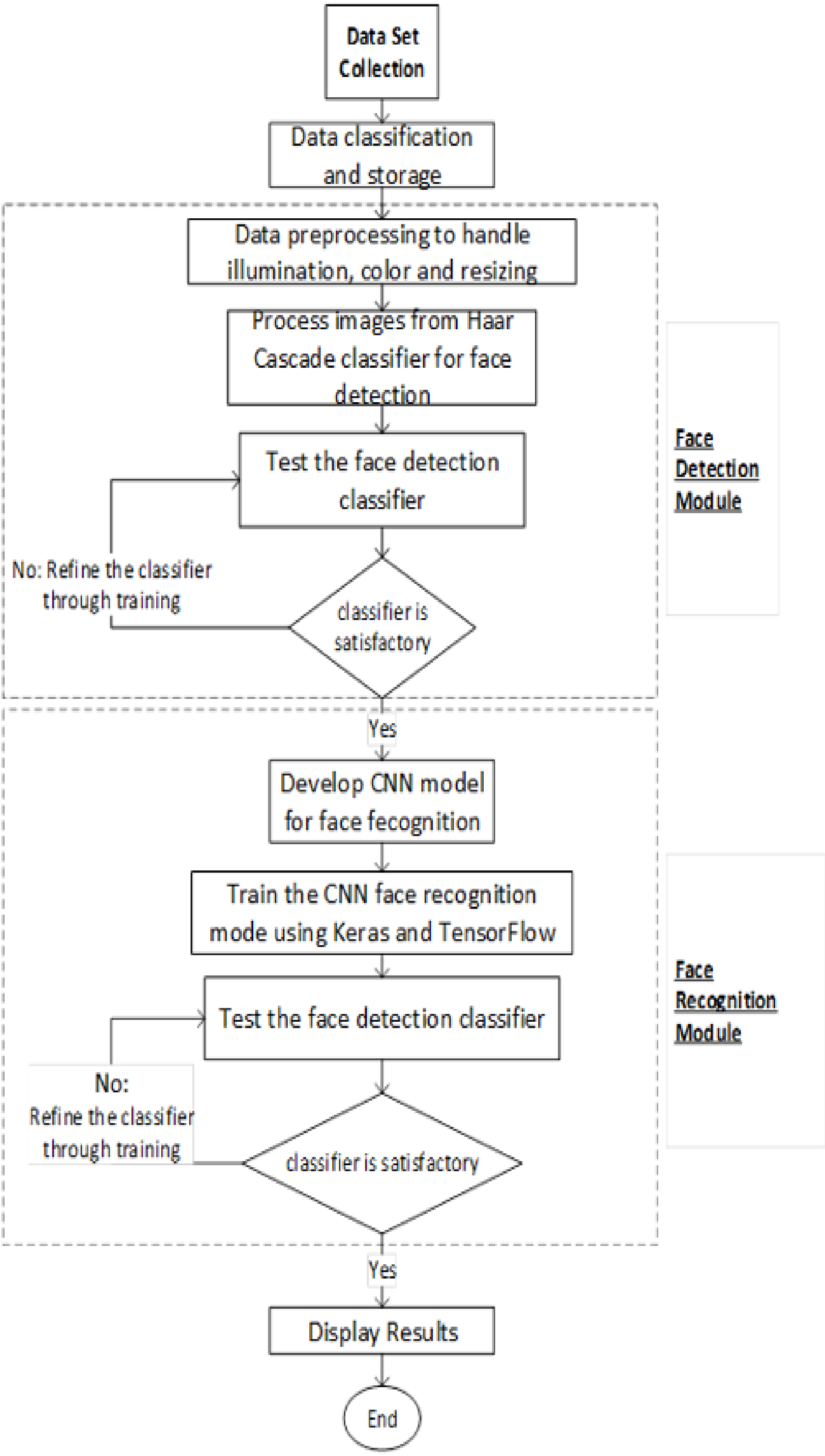


FIGURE 3.5: Combined workflow of face detection and face recognition modules

3.3 Data Execution Flow

The data distribution is made up in three ways:

1. Positive Images:
   This is the first data container contains the same person images which face is required to match.
2. Anchor Images
   This container contain the realtime image of face who is required to verify.
3. Outlier/Negative Images
   This container contains the other faces than the positive and anchor images. But positive image can have positive images too for training purpose.

As shown in is figure 3.5, the data execution flow of a proposed approach for face recognition involves several essential steps that enable the network to learn how to recognize faces accurately. The network takes pairs of face images as input data, with each image going through one of the identical sub-networks. The sub-networks extract a set of features from each input image that captures important visual characteristics such as texture, shape, and edges. These features are then compared to determine the similarity between the two input images.

Next, a loss function is used to calculate the difference between the predicted similarity and the true similarity of the input image pair. The goal is to minimize this difference during training, which leads to a better accuracy of the model. The loss is back propagated through the network to update the weights of each sub-network. This process is repeated for multiple pairs of input images, allowing the network to learn how to extract discriminative features and compare them accurately.

Finally, after training, the proposed network is used for testing by inputting pairs of images that are not used during training. The network is compare the features extracted from each sub-network and output a similarity score between the two images. Overall, the proposed approach for face recognition involves learning a mapping function from pairs of input images to a similarity score using a deep neural network. The network is trained to extract meaningful features from input images and compare them accurately to recognize the identity of the face. This approach is highly effective in recognizing faces even in situations where there are changes in lighting, pose, or facial expression.

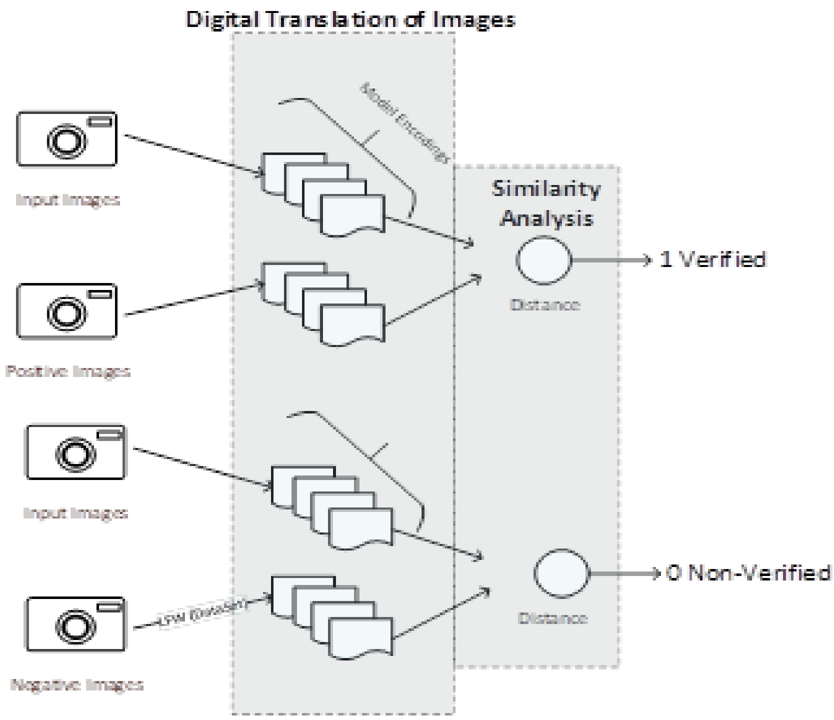


FIGURE 3.6: Data Modeling and Pipelining

3.4 Data Set

The data set is taken from http://vis-www.cs.umass.edu/lfw/. The LFW dataset is abbreviation of "Labeled Faces in the Wild". For the purpose of researching the issue of unrestricted face recognition, these Labeled Faces were photographed in the environment and added to a database. More than 13,000 facial photos gathered from the internet are included in the data collection. Names of the people shown have been written on each face. There are two or more unique photos of 1680 of the individuals included in the data set.

The original and three distinct kinds of "aligned" photographs are now included in four separate sets of LFW images. The aligned pictures consist of "deep filtered" images, "funnelled images" (ICCV 2007), and LFW-a, which employs an undisclosed method of alignment (NIPS 2012). For the majority of face verification algorithms, LFW-a and the deep funnelled photos outperform the original photographs and the funnelled images (ICCV 2007).

The figure 3.6 shows the steps of using the data set and module training. The data set is placed in third container names "Outlier Images". There images of 70% is used to train the model whereas 30% data is used for testing and validation.

The selection of training EPOCS is based on the availability of resources. In particular, where the GPU of the system decides how many EPOCS is applied concurrently to train the model. The Generation 10 computer with the fast GPU is used to perform training which provided the efficiency of 20 EPOCS in less than 1.45 hours.

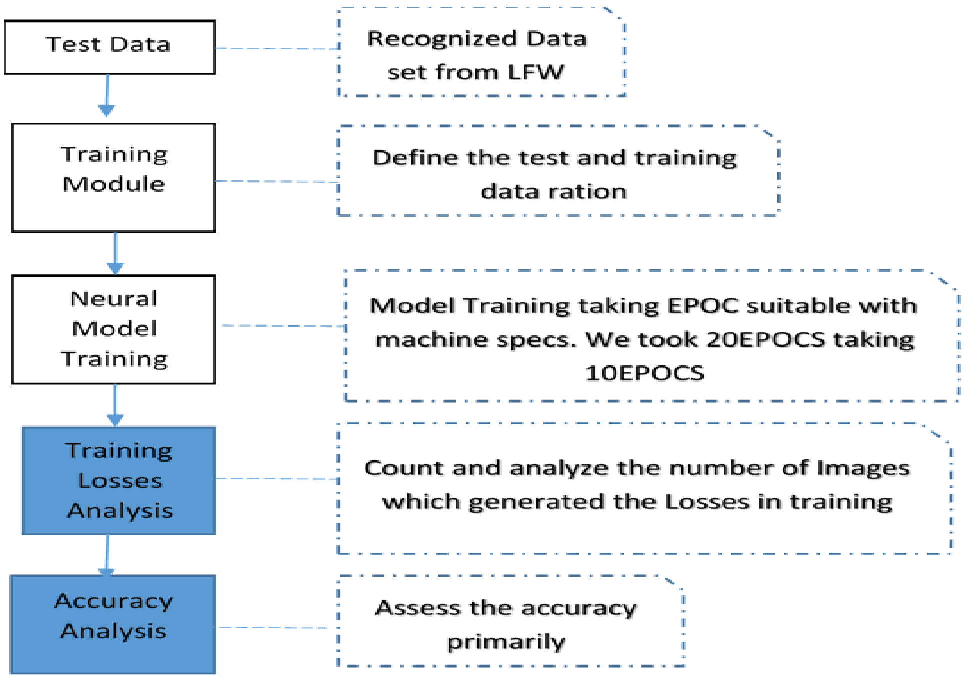


FIGURE 3.7: Training Data Set utilization Steps

The analysis of the image losses is done when training is completed with loss of not more than 3-5%. And a minimum accuracy level of 97% is achieved as planned.

Three image containers is constructed to organize the photographs into categories. One is the real-time image to validate, the “input image”. The positive photographs of the present individual is placed in the second container, which is called ”Positive,” in a variety of settings and positions. Now, photos from both containers is used to confirm the fundamental matching building block. For the creation of the machine-readable encodings, both pictures undergo in preprocessing. These encodings are used to perform input layer preprocessing.

The function is defined to process parallel processing of both images. The preprocessing cycle is take the digital image and settle the illumination of both image streams. The augmentation process is also applied to match the positive containers with Anchor images to achieve the better score of right positive matches.

The difficulty level of provisioned images is kept to train neural network effectively as shown in figure Chapter 3.4. One face picture is captured in different moods and position. Where glasses are also used and hair styling changed. The lighting mode and exposure level also kept different while making photograph of each face. The environment of capture the photograph also kept different while photography. Samples are given below:

The preprocessed images are assigned to distance layer where real time anchor image validated comparing the positive images. The threshold is assigned of 0.5% to give result as 1 as the image is validated whereas the 0 is assigned in case the distance layers gives below score than the 0.5%.

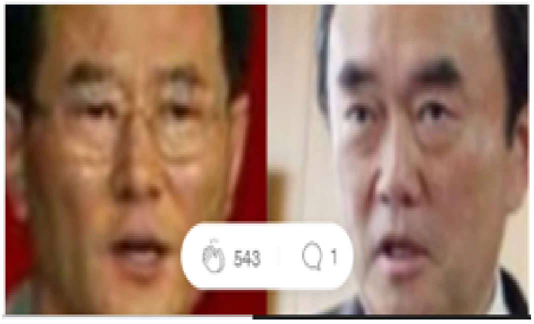
The neural network is designed in a way that is perform given image digital numbers to validate the likelihood among each other through calculating the distance points. The designed neural network is capable of following features:

1. Capable to get trained with common image landscapes and can be utilized for building calculations/predictions about the anonymous class distributions or if few of training data points are available but still workable to make good predictions

Table 3.1: Data Set Complexity Level of Images

(a)

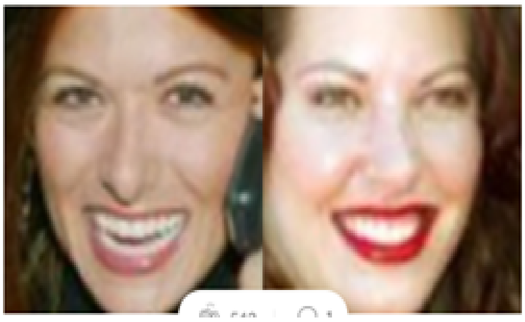
(b)



(c)

(d)

(e)



(f)

(g)

(h)

2. The designed model is capable to train with any data set but we have used the data set_of LFW

Table 3.2: LFW Data Set Samples

3. The deployed neural network is capable providing the competitive approach and enabled to train in deep learning techniques.

4. The neural network is capable of discriminating the class identity of image data points in image verification and recognition process.

The model is hypothesized that this is performed and process the one shot images only. The learning process is adopted to identify the input pairs with the likelihood that the processed image may belongs to same of different classes. But the verification process is based on pair based instead.

While processing both pairing inputs, the pairing with the highest score or the limit of threshold defined for assigning 0 or 1 is marked to get the output of verification or non-verification. In case of positive input pairs the value 1 is assigned to output layer by neural network. Otherwise the 0 value is depict the non-verified images.

3.5 Model Workflow

Figure 3.8 explains the workflow of proposed neural network which is achieved through successful implementation of neural network model along to captures frames at set intervals and perform recognitions to compare the generated encodings.

The developed flow capture the frame from camera and detect human face. The input images come from either still photos or the cut of a broadcast. The image is used as the input for preprocessing, during which the face is identified and a crop provided for the recognition process. To check if the saved picture matches the input image, a comparison is made. Facial points is created for both picture faces. They facial spots are known as the encodings generated where the metrics are applied to the numbers. The difference between the two photographs is then be computed, with the degree of the difference determined to determine if the picture is matching or not. To assure accuracy, the examination of the differences between data encodings is conducted. If false positive findings are achieved, the neural network is fed additional data as a training set or the neural network is trained using data points of that image's type. If the desired results are not obtained, this process is repeated.

The proposed approach of this study applied where two images are processed to verify the recognition output. The neural network built on adopting the CNN layered approach where multiple layer contributing to perform the verification task.

The input layer takes two images and process parallel to translate into machine readable format. The preprocessing is performed and encodings are generated of both images. The 2Diementional encodings is processed at Distance layer. Both images encodings are validating and the difference of set threshold is give result of matching and non-matching of both images.

The proposed model is able to integrate with the existing system. The input layer needs the still image or video stream frame collected for preprocessing, and then the ovulational layer and Rectified linear unit (ReLu) execute convolution operations on the elements to create the feature map. When doing matching on the resulting encoding, maxpooling is passing the two-dimensional array. The results of the matching compare the established threshold with the discovered matching. If the threshold satisfies the required standards, the results are shown as picture verified, and vice versa. The source images might be either static pictures or a broadcast's edit. The image is the input for the preprocessing, which is identify the face and offer a crop for the identification procedure. A comparison is performed to see if the stored image matches the input image. For both photo faces, facial points is made. The metrics is applied to the numbers in these face regions, which are referred to as the encodings.

The degree of the discrepancy between the two photos is then be calculated in order to establish whether or not the picture matches. Examining the variations in data encodings is done in order to guarantee correctness. If false positive results are obtained, the neural network is either be trained using data points specific to that sort of picture or is fed more data as a training set. This procedure is repeated if the intended outcomes are not attained the first time.

As shown in figure 3.8, the proposed model is a type of neural network that is used for tasks that involve determining the similarity between two input images. This is used to determine whether two images belong to the same person or not.

The architecture of the proposed network consists of two identical sub-networks that are designed to process each input image independently. These sub-networks are typically convolutional neural networks (CNNs) that are used to extract features from the input images.

The input images are passed through each sub-network, and the output of each sub-network is then compared to determine the similarity between the two images. The comparison is typically done using a distance metric. The output of the proposed network is a similarity score between 0 and 1, where 0 indicates that the two input images are completely dissimilar, and 1 indicates that they are identical.

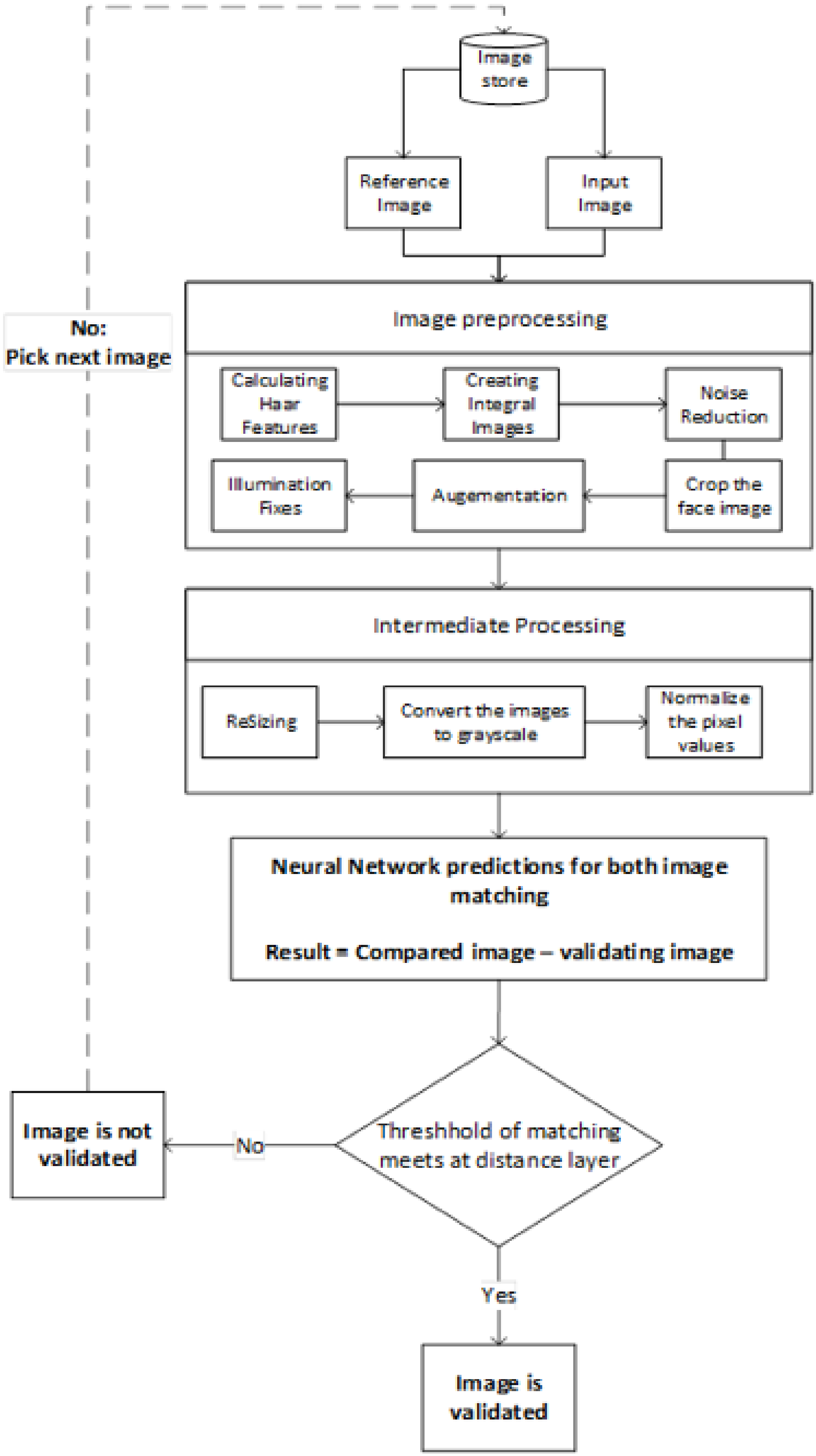


FIGURE 3.8: Workflow of Face Recognition System

This similarity score is used to determine whether two input images belong to the same person or not.

The training process for the Proposed network involves using pairs of images as input and a target output value that indicates whether the two images belong to the same person or not. The network is trained using a loss function that encourages the network to output a high similarity score for pairs of images that belong to the same person and a low similarity score for pairs of images that belong to different people.

Once the Proposed network is trained, it is used to perform face recognition tasks by comparing pairs of input images and determining their similarity score. This is done by extracting features from each image using the sub-networks and then computing the similarity score using the distance metric.

3.6 Pseudo Code of Proposed Neural Network

Ensure: train.CNN_Network = CNet

Ensure: train.FDHAAR_Network = FDNet

Ensure: Dataset.LFQ = Augmentation

Ensure: DataSplit = Positive , Negative, InputImage

To deploy Face detection module

Download : HAARLikeFeature.xml

Load: Run → Detection Algorithm

Input_image = noise_reduction + Illumination+resizing + cropping

Augmentation = Input_image

Augemented_FacialDatapoints = augmentation

decteted_face_Imgae = Augemented_FacialDatapoints

Face_recognition_module = decteted_face_Imgae

To develop CNN for face recognition

Nerual Layers definition

input = Input(shape=input_image)

x = Dense(128, activation='relu')(input)

x = Dense(64, activation='relu')(x)

x = Dense(32, activation='relu')(x)

return Model(input, x)

Creates the neural model

base_network = create_base_network(input_shape)

input_a = Input(shape=input_shape)

input_b = Input(shape=input_shape)

processed_a = base_network(input_a)

processed_b = base_network(input_b)

distance = ([processed_a, processed_b])

model = Model(inputs=[input_a, input_b], outputs=distance)

return model

Compile the model

optimizer = RMSprop()

Neural_network.compile(loss='binary_crossentropy', optimizer=optimizer)

Train the model

Using Negative or Suspect training data of LFW which was augmented

X_train = [...]

Y_train = [...] # update training labels

Neural_network.fit([X_train[:, 0], X_train[:, 1]], Y_train, epochs=10, batch_size=128)

Test the model

To compare two images

Input_image → frameCapture

Positive_Image → image pool

Negative_image → Suspect pool

from positiveimage.measure import compare_input

def distance(imageA, imageB):

return compare_input(imageA, imageB, multichannel=True) or

return compare_input(imageA, imageB, multichannel=True)

Chapter 4

Performance Evaluation

The proposed neural network is used to compare and match pairs of images to identify similarity. During training, the neural network learns to generate a similarity score between two images that indicates how similar they are. This similarity score is then used to make a binary decision about whether the two images belong to the same category or not.

Once the training of the developed neural network is completed, its performance is evaluated using several evaluation metrics such as accuracy, precision, and recall.

4.1 Performance Evaluation Approaches

The **accuracy approach** is used to evaluate the performance of the neural network in assigning input images to one of several predefined categories. In the context of face recognition, this means that the neural network is trained to identify whether two face images belong to the same person or not. The accuracy metric directly measures the proportion of correctly classified face pairs, which is a key aspect of the task.

In the face recognition module, the accuracy of the proposed neural network is evaluated by calculating the proportion of correctly classified face pairs. A pair is considered to be a match if the similarity score generated by the neural network exceeds a predefined threshold of 0.5. This threshold is used to decide whether the two images are similar enough to be considered a match or not.

To evaluate the accuracy of the proposed network, a manual evaluation is performed on a set of sample face pairs. This evaluation is performed as the True Acceptance Rate (TAR) or True Positive Rate (TPR), which measures the proportion of genuine matches that are correctly recognized as an output. The TAR or TPR is manually calculated by dividing the number of correctly classified face pairs by the total number of genuine matches in the evaluation set.

In addition to accuracy, precision and recall are also used evaluation metrics for proposed neural networks.

The **precision approach** is used as a performance evaluation metric for face recognition proposed neural networks because it measures the accuracy of the network in identifying similar or dissimilar pairs of images.

In a proposed neural network for face recognition, the network is trained to compare two input images and determine whether they are of the same person or not. The precision metric measures the ratio of true positive results (i.e., the network correctly identifies that two images are of the same person) to the total number of positive results (i.e., the network identifies that two images are of the same person regardless of whether this is true or not).

Precision focuses on the accuracy of positive predictions, which is particularly important in face recognition tasks where false positives (i.e., identifying two different people as the same person) can have serious consequences.

The used formula for precision is:

$$precision = \frac{truepositives}{(truepositives + falsepositives)}$$

Where true positives are the number of correctly identified positive instances (i.e., the number of times the proposed neural network correctly identified that two images belong to the same person), and false positives are the number of instances where the network incorrectly identified two different images as belonging to the same person.

Precision formula helped to measure of how precise or accurate the network is in identifying positive instances, which in the context of face recognition proposed neural networks, is the ability to correctly identify whether two images belong to the same person or not.

Recall approach is used performance evaluation metric for face recognition proposed neural networks, along with precision. Recall measures the ability of the network to correctly identify all positive instances (i.e., the number of times the network correctly identified two images as belonging to the same person), irrespective of whether some positive instances were missed by the network or not.

In face recognition module, recall calculated false negatives (i.e., the network failing to identify two images belonging to the same person) can have serious consequences, such as failing to detect a security threat or allowing unauthorized access to a secure facility.

The recall formula is:

$$recall = \frac{truepositives}{(truepositives + falsenegatives)}$$

Where true positives are the number of correctly identified positive instances, and false negatives are the number of instances where the network failed to identify two images belonging to the same person.

While precision measures the accuracy of positive predictions, recall measures the completeness of positive predictions. Both approaches are evaluating the performance of a face recognition proposed neural network.

4.2 Experiment Results

The test data was split for validation and applied to six layers of neural network model. The optimizer is configured to improve the image validation set's accuracy. In order to save computing time, we limited training runs to 50 epochs because by this point in the learning schedule, a good approximation could typically be produced. Although the maximum number of Iterations are made up of 100. Convolutional networks were actually used with fewer iterations since the calculation time was too expensive (several hours to several days). Additionally, we developed visualizations that translate the training and test sets into the lower-dimensional feature spaces used by the network to compute distances and generate predictions. The first histogram demonstrates how the neural network learns to increase or decrease the distance allocated to each training instance based on its label. The second histogram displays the distribution and size of forecasts for each data point in the test set that belongs to either class.

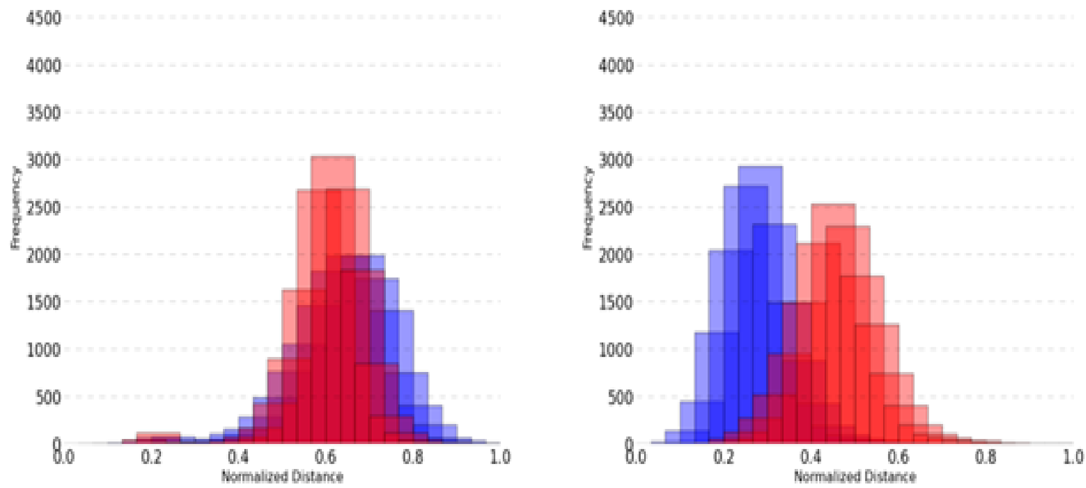


FIGURE 4.1: Distance Layer

Left Histogram (Before Training) describes position of data matching before training.

Right Histogram describes the neural network data point's strength to calculate distances after training.

The difference between the left and right histogram clearly describing the improvement in learning where the neural network is able to identify the distances between images before and after training.

Below table is describing the number of training sets used to bring accuracy in distance measurement between input and matching images.

Table 4.1: Accuracy on verification task (convolutional neural net)

Training Sets	**Accuracy on Test (6-layer)**
30k training	
no distortions	90.61
affine distortions x8	91.9
60k training	
no distortions	91.54
affine distortions x8	93.15
90k training	
no distortions	91.63
affine distortions x8	93.42

The verification network was successful in both achieving high accuracy and displaying a high level of confidence in its predictions.

The entire embedded layers are prepared where the input and validation image are processed to analyses the function output. The network layers are embedded to review the output as given below:

```
siamese_network.summary()

Model: "SiameseNetwork"
__________________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_img (InputLayer)         [(None, 100, 100, 3  0           []
                                )]

 validation_img (InputLayer)    [(None, 100, 100, 3  0           []
                                )]

 embedding (Functional)         (None, 4096)         38960448    ['input_img[0][0]',
                                                                  'validation_img[0][0]']

 l1_dist_1 (L1Dist)             (None, 4096)         0           ['embedding[0][0]',
                                                                  'embedding[1][0]']

 dense_2 (Dense)                (None, 1)            4097        ['l1_dist_1[0][0]']

==================================================================================================
Total params: 38,964,545
Trainable params: 38,964,545
Non-trainable params: 0
__________________________________________________________________________________________________
```

FIGURE 4.2: Proposed neural network layers Summary

4.2.1 Precision Results

Precision approach is applied for evaluation to verify the proportion of positive identifications made manually were actually corrected. The formula is applied to calculate the precision which is explained in detail in performance evaluation approaches to evaluate the neural network performance evaluation.

The precision results was received as 0 means that the input images gave correct results of matching the input and matching images and didn't produce the false positive results. Where false positive describe the images which were not matching but system has shown in positive matchings. In our precision results the false positive rate remain 0. It means that the false negative remained identified as not matching while processing the similarities and identifying differences.

4.2.2 Recall Results

Recall attempts to answer about the proportion of actual positives identified by classifier were identified correctly. Which remain 1, describing that the input images which should be matches actually identifies as matched by the proposed neural network. The below figure is taken from the code results where the recall function was called to describe the positive matchings.

```
# Creating a metric object
m = Precision()

# Calculating the recall value
m.update_state(y_true, y_hat)

# Return Recall Result
m.result().numpy()

1.0
```

FIGURE 4.3: Recall results of proposed neural network

4.2.3 Accuracy Results

The identified combinations of True and False Positives and Negatives were closely assessed in precision and recall. Where the accuracy of the model was checked technically by using following formula:-

$$\frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositives + FalseNegatives}$$

$$= \frac{N.ofCorrectPredictions}{N.ofallPredictions} = \frac{N.ofCorrectPredictions}{SizeofDataset}$$

The EPOCH are set to the 50 with batch of 16 to train the 300 images for positive and input image data set whereas the negative dataset is prepared with 70% more of this number. The accuracy is achieved at 100% of the model got trained with the data set.

The model training was performed in phases where data was fed in batches to check the accuracy. It was found that initial batches of data in anchor and outlier images could bring the low accuracy till 40%. Then the data set was increased and augmentation was applied in 50,000 images sets.

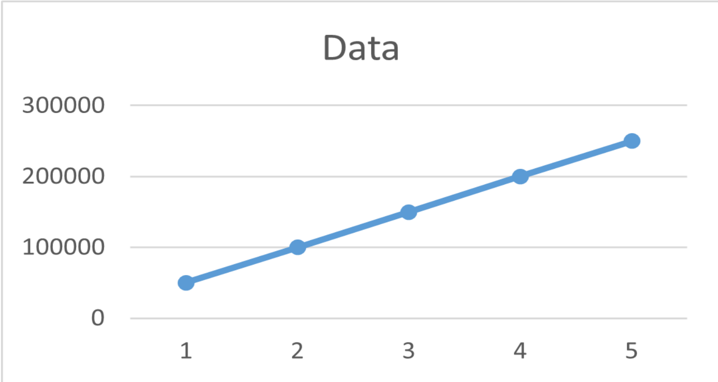


FIGURE 4.4: Data batches for Training

The data feed improved the accuracy till 98% and above. It was observed that feeding the more data was not brining the improvement but only the resource utilization were increased. So the data set was only fed till 300,000 achieving the 98% of accuracy.

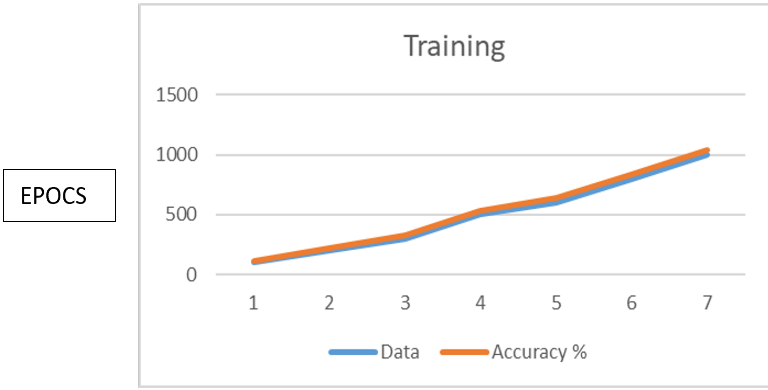


FIGURE 4.5: Training of Module in Data Batches to Achieve Accuracy

The Augmentation process was performed on image visualization improvement. This process was used to improve the Contract, Resizing, Pose improvement and Brightness which has increased the data set as well as the data improvement. Which brought the proposed neural network on highest score of accuracy.

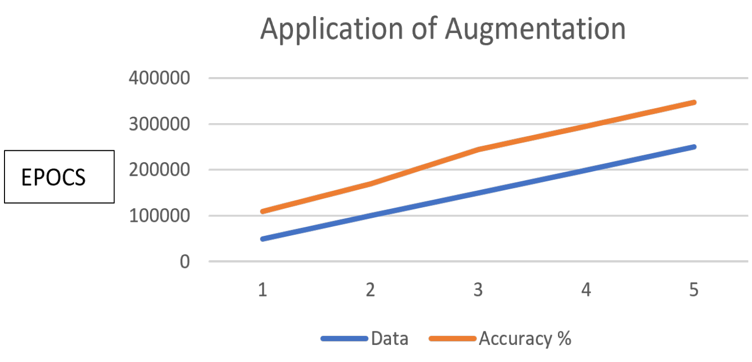


FIGURE 4.6: Data Augmentation to improve the Accuracy

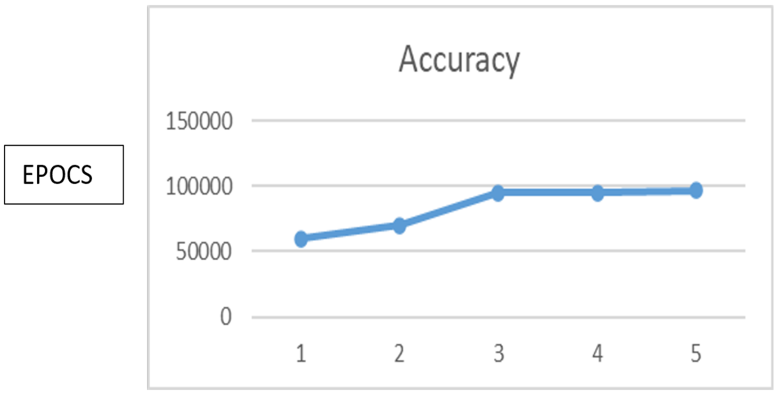


FIGURE 4.7: Accuracy Curve in Training

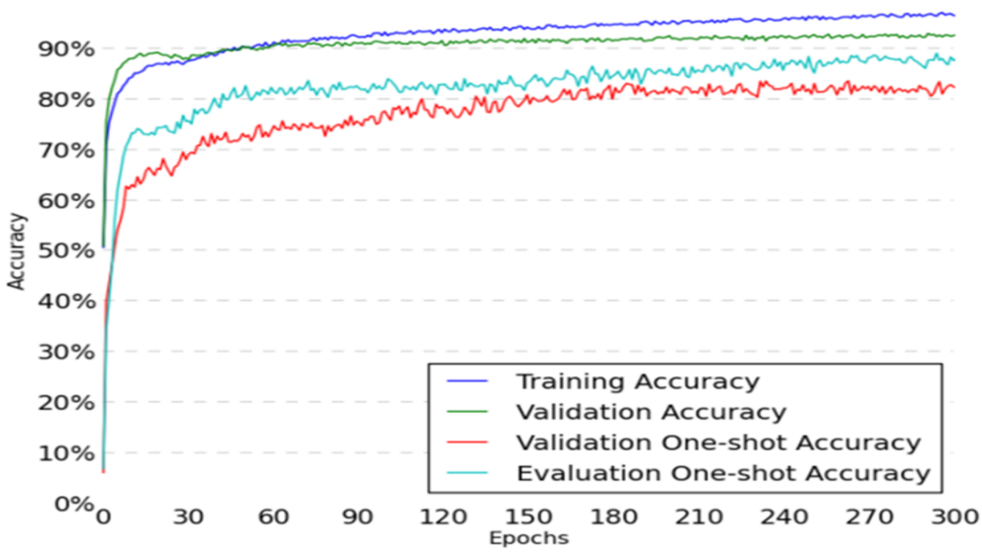


FIGURE 4.8: Model Accuracy with Training

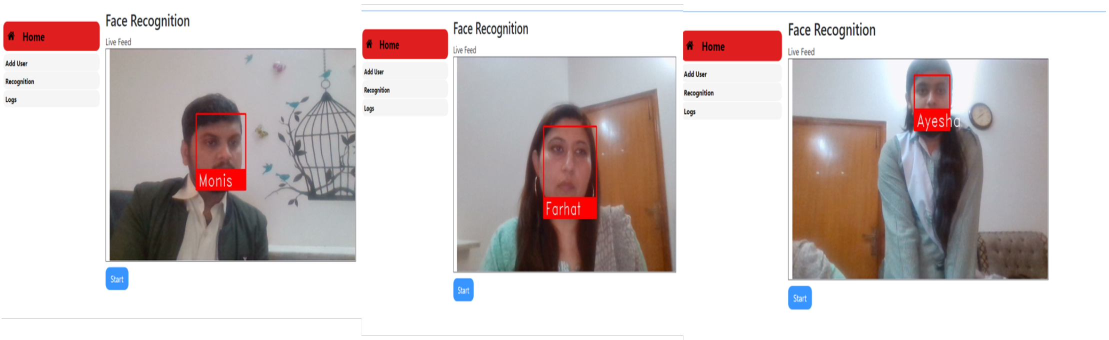


FIGURE 4.9: Recognition Accuracy of Second Record

4.3 Comparison with Other Models

The accuracy of multiple model is compared with other neural models. The pre-trained xml files were loaded into the system where the detection module gave the output to the face recognition system using HOG algorithm. The HOG system gave results less than the results provided by proposed system. The same approach was used to evaluate the PCA where the results accuracy remain less than the proposed model.

The proposed model went on highest accuracy like 98%. Whereas the HOG and PCA remained 96% and 97% respectively.

FIGURE 4.10: Comparison Chart of Accuracy

Chapter 5

Conclusions and Future Work

5.1 Conclusions and Future Work

This study comprises of two area of domains. First is to explore the neural network for face recognition technology which should be able to perform face detection and recognition with accuracy. Second is to identify the appropriate process of border control where the neural network can implement within existing system or deploy as independent system. The overall objective of this study is to develop the facial recognition neural network with improved learning which can help to identify suspect face from processed images as an input and the image used to compare. This neural network also able to handle face recognition open and closed environment challenges e.g. pose, illumination, image quality and sizing etc.

The first research question was established as how can an appropriate data set be used to train a facial recognition neural network which can resolve illumination, pose and environment challenges to perform human features recognition?. The face detection algorithms include HAAR feature selection, creating integral image, AdaBoosting and casecading classifiers were studied and compared the results. The HAAR feature selection is chosen to perform the face detection process. Whereas the face recognition algorithms were studied includes eigenfaces, fisherfaces, local binary patterns and Convolutional neural networks (CNN) and compared the technological capabilities. The CNN found most suitable to use for this study. The neural network is developed with six layers of neurons and trained using world recognized LFW dataset for facial recognition. The diversity of data points is enhanced using augmentation process with pose, resizing, brightness and resolution improvements. The results are evaluated using accuracy, precision and recall methods and performed backpropagation to achieve maximum efficiency. The proposed results was able to produce the 98% of accuracy while testing the results applying close to real scenarios of border control open and closed environment. The results were compared with HOG and PCA which were producing maximum 95.5-97% results with low precision and recall than the proposed neural network.

The second research question was as how can facial recognition technology be effectively integrated into the border control management process? The border control processes were studied of multiple countries. The process areas were studied and identified Traveler profile verification process appropriate to integrate the proposed neural network. The process of traveler authentication is further studied with 1:1 and 1:N perspective. Where 1:1 is the subprocess where the national data is already retained by the country and can be verified by putting the identification number e.g NADRA CNIC is used in Pakistan. The other 1:N sub process is studied where the suspect data is retrained as negative pool means that the stored faces should not cross boarder for In and Out both. This pool includes foreigners and national suspect faces. The proposed model is tested for both sub processes including the illumination, pose, sizing and resolution diversity and found 98

The proposed solution is found appropriate to implement at border control to manage and control the authenticate traffic of individuals. Our proposed model and point of integration found the most appropriate to embed in existing border control Process.

The proposed solution is also capable to perform image verification with real time incoming feed with high level of accuracy. Other institutes intending to authenticate their human traffic can use the proposed solution as well. The strength of this system is that it takes few images to gather the data points and gives highly accurate results in response.

The technique should be extended to improve the speed with accuracy by training the neural model. The facial recognition system processing the personal information which is at high risks of data security. The whole processing of information is using application and network controls. The multiple components of technology are utilized to produce results. This is important to work of technical security gaps to resolve. Future work may include branching out to earning the system with different poses like at 90 Degree angle and dealing the data security challenges.

Bibliography

[1] D. Balya and T. Roska, "Face and eye detection by cnn algorithms," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 23, pp. 497–511, 1999.

[2] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern recognition*, vol. 27, no. 1, pp. 53–63, 1994.

[3] Mosaic, "Mosaic faces images - search images on everypixel." https://www.everypixel.com/q/mosaic-faces?image_id=4025919532289504429, March 2023. (Accessed on 04/11/2023).

[4] L. R. Carlos-Roca, I. H. Torres, and C. F. Tena, "Facial recognition application for border control," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.

[5] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2537–2540, IEEE, 1997.

[6] J. A. Shufelt, "Performance evaluation and analysis of monocular building extraction from aerial imagery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 311–326, 1999.

[7] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[8] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[9] T. Sakai, M. Nagao, and S. Fujibayashi, "Line extraction and pattern detection in a photograph," *Pattern recognition*, vol. 1, no. 3, pp. 233–248, 1969.

[10] I. Craw, H. Ellis, and J. Lishman, "Automatic extraction of face-features," *Pattern Recognition Letters*, vol. 5, no. 2, pp. 183–187, 1987.

[11] G. Hu, Y. Yang, D. Yi, J. Kittler, W. Christmas, S. Z. Li, and T. Hospedales, "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition," in *Proceedings of the IEEE international conference on computer vision workshops*, pp. 142–150, 2015.

[12] L. R. Carlos-Roca, I. H. Torres, and C. F. Tena, "Facial recognition application for border control," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.

[13] R. Patel and S. Patel, "A comprehensive study of applying convolutional neural network for computer vision," *International Journal of Advanced Science and Technology*, vol. 6, no. 6, pp. 2161–2174, 2020.

[14] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*, pp. 499–515, Springer, 2016.

[15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[16] A. N. Razzaq, R. Ghazali, N. K. El Abbadi, and M. Dosh, "A comprehensive survey on face detection techniques," *Webology*, vol. 19, no. 1, pp. 613–628, 2022.

[17] C. Rathgeb, A. Dantcheva, and C. Busch, "Impact and detection of facial beautification in face recognition: An overview," *IEEE Access*, vol. 7, pp. 152667–152678, 2019.

[18] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[19] E. Garage, "Face recognition." https://www.engineersgarage.com/face-recognition/, March 2023. (Accessed on 04/11/2023).

[20] M. Khan, S. Chakraborty, R. Astya, and S. Khepra, "Face detection and recognition using opencv," in *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 116–119, IEEE, 2019.

[21] Wikipedia, "Photometric stereo - wikipedia." https://en.wikipedia.org/wiki/Photometric_stereo, March 2023. (Accessed on 04/11/2023).

[22] Roseman, "Video streaming services roadmap for content suppliers, and enhanced portals." https://www.rosemansolutions.com/wordpress/2014/10/video-streaming-services-roadmap-content-suppliers-enhanced-portals/, March 2023. (Accessed on 04/11/2023).

[23] M. üge Çarıkçı and F. Özen, "A face recognition system based on eigenfaces method," *Procedia Technology*, vol. 1, pp. 118–123, 2012.

[24] L. R. Carlos-Roca, I. H. Torres, and C. F. Tena, "Facial recognition application for border control," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.

[25] C. Bhargav and T. Chakrapani, "A review on image processing," *computer*, vol. 5, no. 02, p. 2018.

[26] L. Sirovich, M. Kirby, and M. Winter, "An eigenfunction approach to large scale transitional structures in jet flow," *Physics of Fluids A: Fluid Dynamics*, vol. 2, no. 2, pp. 127–136, 1990.

[27] L. R. Carlos-Roca, I. H. Torres, and C. F. Tena, "Facial recognition application for border control," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.

[28] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[29] B. Kocacinar, B. Tas, F. P. Akbulut, C. Catal, and D. Mishra, "A real-time cnn-based lightweight mobile masked face recognition system," *Ieee Access*, vol. 10, pp. 63496–63507, 2022.

[30] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137–154, 2004.

[31] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580–587, 2014.

[32] L. Li, X. Mu, S. Li, and H. Peng, "A review of face recognition technology," *IEEE access*, vol. 8, pp. 139110–139120, 2020.

[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," 2015.

[34] R. Meriton, R. Bhandal, G. Graham, and A. Brown, "An examination of the generative mechanisms of value in big data-enabled supply chain management research," *International Journal of Production Research*, vol. 59, no. 23, pp. 7283–7310, 2021.

[35] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 886–893, Ieee, 2005.

[36] M. García-Gómez, M. Antonio-Cruz, and A. Y. Castañeda-Miyamoto, "A review on face recognition-based access control systems," in *2021 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, vol. 5, pp. 1–11, IEEE, 2021.

[37] E. Pass, "Bpol - homepage - http status 404." https://www.easypass.de/EasyPass/EN/EasyPASS%20RTP/rtp_node.html, March 2023. (Accessed on 04/11/2023).

[38] W. Zhao and R. Chellappa, *Face Processing: Advanced modeling and methods.* Elsevier, 2011.

[39] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 10, pp. 1042–1052, 1993.

[40] N. K. Singh and K. Yadav, "A novel approach for developed face recognition system using neural network," *Think India Journal*, vol. 21, no. 1, pp. 62–71, 2018.

[41] IATA, "Iata - 404 page." https://www.iata.org/en/programs/promoting-aviation, March 2023. (Accessed on 04/11/2023).

[42] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1701–1708, 2014.

[43] U. Customs, "Requested page not found — u.s. customs and border protection." https://www.cbp.gov/travel/us%20citizens/apc, March 2023. (Accessed on 04/11/2023).

[44] N. Khan and M. Efthymiou, "The use of biometric technology at airports: The case of customs and border protection (cbp)," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100049, 2021.

[45] Arxiv, "Arxiv." https://arxiv.org/pdf/1904.11685.pdf, March 2023. (Accessed on 04/12/2023).