Facial Recognition Using AI

# ABSTRACT

The automatic face recognition attendance system was designed with the aim of marking attendance of the students present in a classroom based on facial recognition and give out a marked attendance sheet.

The system provides an efficient way of marking and storing the attendance without having to physically call out the name of each student. It helps save time and the attendance shall be directly stored without having to maintain a physical record.

# General Terms

Face recognition, face detection, eigen-faces, eigen-vectors, servo motors, Arduino micro-controller, openCV, auto-focus, Laplace transform, webcam, Principle Component Analysis (PCA).

# INTRODUCTION

# Faces are amongst the most relevant social stimuli as they communicate information essential for the course of social interaction and communication. Specifically, facial expressions convey information about what emotion is currently experienced by a target, which in turn affects how the target is perceived and what behavioral tendencies are elicited in the observer.

# n this vein, the present research links facial emotion recognition to broad and basic motivational orientations, namely regulatory focus. Studying the impact of regulatory focus on facial emotion recognition seemed to us particularly worthwhile, because regulatory focus introduces a fundamental distinction between two motivational strategies that have been shown to affect an enormous range of phenomena ranging from basic motivational mechanisms to intergroup relations.

# Motivation

# Facial recognition technology has gained significant attention and adoption in recent years due to several key motivations. Here are some of the primary motivations behind the development and use of facial recognition:

# 1) Enhanced Security: Facial recognition offers improved security by enabling accurate and efficient identification of individuals. It can be employed in various scenarios such as airport security, access control systems, and law

# enforcement to help prevent unauthorized access, identify potential threats, and assist in criminal investigations.

# 2) Convenient Authentication: Facial recognition provides a convenient and contactless method of authentication.

# Background

# Facial recognition is more than 50 years old.

# A research team led by Woodrow W Bledsoe ran experiments between 1964 and 1966 to see whether ‘programming computers’ could recognize human faces.

# Bledsoe said: “The face recognition problem is made difficult by the great variability in head rotation and tilt, lighting intensity and angle, facial expression, aging, etc.”

# Facial recognition has come of age thanks to camera technology improvements, mapping processes, machine learning, and processing speeds.

# 1.3 Objective

# The objective of face recognition is, from the incoming image, to find a series of data of the same face in a set of training images in a database. The great difficulty is ensuring that this process is carried out in real-time, something that is not available to all biometric face recognition software providers.

# These compare the relevant information of the incoming image signal in real-time in a photo or video in a database, being much more reliable and secure than the information obtained in a static image. This biometric facial recognition procedure requires an internet connection since the database cannot be located on the capture device as it is hosted on servers.

# Fundamentals of Face Recognition

# It compares the person's face to the one in the profile present in the database to check if they match. It's a 1-to-1 matching system as the system has to match the individual's face against a specific face already present in the linked profile. Thus, verification is quicker than identification and more accurate.

# 2.1 Face Representation

# Human faces have to be represented before recognition can take place. The role played by representation is most important, and it probably exceeds that played by recognition, known also as classification or identification.

# Kant described the problem of how anything in the mind can be a “representation” of anything outside the mind as the most difficult riddle in philosophy.

# We can recognize transformed and dimensionally-reduced as well as original human faces in their “raw” form. Gerald Edelman (1987) has described innate neural architectures that generically implement what is referred to as neural Darwinism.

# The non-accidental properties of human faces are the particular dimensions the human face representations become tuned for. Such properties eventually induce the features extracted to represent human faces. The features correspond to coordinate axes and they define the face space.

# It corresponds to closed loop systems whose behavior is driven by performance.

Features are proposed ("extracted") and kept ("selected") according to their utility in recognition. . Similar viewpoints have also been formulated by Barlow

(1989a,b)

# (a) Statistical characterization of natural images

(Ruder man, 1994); andand

(b) relationships between the statistical properties of natural Images and the optimization of the visual system. Specific examples of such optimization

Criteria and their realization for representing human faces include distillation and minimization of the reconstruction error leading to principal component analysis.

**2.1.1** **Face shape**

In facial recognition technology, face shape is one of the key factors used to identify and classify faces. AI-based facial recognition systems analyze the geometric structure of a face to determine its shape and extract features that are unique to that individual. Here's a general overview of how face shape is utilized in facial recognition:

1. Face Detection: The first step in facial recognition is detecting and locating a face within an image or video frame. Algorithms scan the input data, identify regions that potentially contain a face, and create

bounding boxes around them.

1. Feature Extraction: After locating the facial landmarks, the system extracts various facial features from the face, including the shape of the jawline, the distance between the eyes, the width of the nose, etc. These features are then used to create a unique representation, often referred to as a face template or face embedding, which captures the distinct characteristics of the individual's face.
2. Face Matching: In the identification or verification phase, the face template extracted from the input face is compared against a database of known faces. The system calculates the similarity or distance between the templates to determine if there's a match or a potential identity.
3. Facial Landmark Detection: Once a face is detected, facial landmark detection algorithms are employed to locate specific points on the face, such as the corners of the eyes, nose, and mouth. These landmarks provide crucial reference points for analyzing the face's shape and features.
4. Machine Learning and Training: To recognize faces accurately, AI models are trained on large datasets containing labeled images of faces. These models learn patterns and correlations between facial features and face shape through deep learning techniques, such as convolutional neural networks (CNNs) and embeddings.

**2.2 Feature Extraction**

Since the last century biometric techniques were used for identification of humans. Faces are one of many forms of biometrics used to identify individuals and to verify their identity. Face recognition refers to the automated method of verifying a match between two human faces. Feature extraction is a very important step in face recognition. The recognition rate of system depends on the meaningful data extracted from the face image. If the features belong to different classes and the distance between these classes is large then these features are important for a given image. There is no 100% matching between the images of the same face even if they were from the same person. In this study, the analysis of face recognition systems using **three different feature extraction techniques:-**

1. **Principal Component Analysis (PCA),**
2. **Fisher Linear Discriminant analysis (FLD)**
3. **Fast Pixel Based Matching (FPBM)**
4. **Principal Component Analysis (PCA)**

PCA is a technique that takes high-dimensional image data and uses the dependencies between the variables to represent it in a more tractable, lower dimensional form, without losing too much information. PCA is a statistical procedure that evaluates the covariance structure of a set of variables and identifies the principal directions in data variables. PCA is used to identify sets of orthogonal coordinate axes through the data. Principal components are determined by computing eigenvectors and eigenvalues of the data covariance matrix. Based on principal components the identification of face images is performed.

1. **Fisher Linear Discriminant analysis (FLD)**

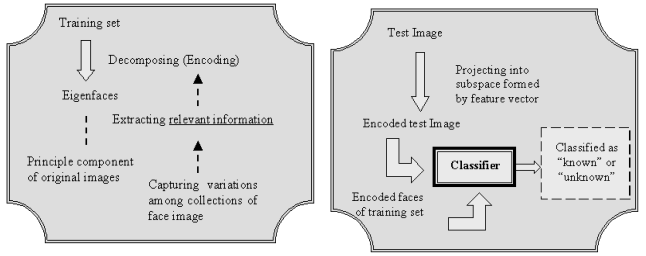
FLD is the most famous way to search for trends in the data, which has the largest difference and highlight data. This method is also used, for lower-dimensional representation of the data, which removes some of the trends “noisy”. The basic idea of FLDA is the design an optimal transform, which can maximize the ratio of between-class to within-class scatter matrices so that the classes can be well separated in the low-dimensional space. FLD method allows information between members of the same category (images of the same person) to develop a set of feature vectors. FLD uses a linear projection of the n-dimensional data onto a one-dimensional space (i.e., line). Projection onto a line is separated by a class and classification problem becomes choosing a line.

1. **Fast Pixel Based Matching (FPBM)**

FPBM is a method to extract the features of the images on the basis of matching image areas and sub pixel displacement estimate using similarity measures. The recognition is based on the edge detection. This method generates much less information than the original image has. This is because it eliminates most of the details that are not relevant for the purpose of identifying the boundaries, while preserving the essential information to describe the shape and structural characteristics and geometry of the objects represented. This study describes the design of a face recognition system using PCA, FLD and FPBM methods. Each of these techniques was implemented in MATLAB. The outputs of the feature extraction block are classified to recognize the face patterns. The algorithm uses Euclidean Distance for classification of face images. Comparisons of the simulation results of face recognition systems using PCA, FLD and FPBM algorithms are presented.

**2.2.1 Principal Component Analysis**

PCA is based on an information theory approach that decomposes face images into small set of feature images called “Eigenfaces” which may be thought of as a principle component analysis of original training set of face images. In order to decompose, I have to extract relevant information from face image. A simple approach is to capture the variations from a collection of training face images, independent of any judgment of features (i.e. second-order statistics of the data are de-correlated) and use this information to encode and compare individuals as shown in Fig. 3a and 3b.

**a)**.Decomposing the training **b)** Classification of the input

face image face images

The algorithm is as follows:

Step 1:

**Establishes the training set**



Step 2:

**Calculate mean image of all training samples.**



Where is a mean, M is number of faces in database

Step 3:

Calculates the difference Images by subtracting the training set vector by the mean image. Let us call this matrix as the variation matrix.

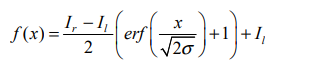


Thus, the egan values are found by and can be further used to calculate eganfaces.

2.2.4 **Fast Pixel Based Matching using Edge Detection**

The recognition of the contours (edge detection) is used for the purpose of marking the points of a digital image in which the light intensity changes abruptly. Abrupt changes of the properties of an image are usually a symptom of events or major changes of the physical world. These changes can be discontinuity of the depth in the surface, changing the properties of materials and variations in lighting conditions from the surrounding environment. The edge detection is a research field of image processing, particularly the branch of feature recognition. The operation of edge detection generates images containing much less information than the original, because it eliminates most of the details that are not relevant for the purpose of identifying the boundaries (Marr and Hildreth, 1980; Fraser, 1985). The methods used for edge detections can be grouped into two categories: Search-based and zero-crossing based. Search based methods recognize the contours by computing the maxima and the minima of the first order derivative of the image, usually looking in the direction in which we have the maximum local gradient. Zero-crossing based methods seek for the zero-crossing points at which the derivative of the second order passes through zero, usually the Laplacian function or a differential expression of a non-linear function.

The contours play a very important role in many applications of computer vision. A typical contour could be, for example, the boundary between an area of red colour and a yellow, or a line with a thickness of a few pixels and a different colour compared to a uniform colour background. The model illustrated here, has a function error erf that can be used to create a mathematical model of the effects of the blurs sufficient accurate to describe many practical applications. An image f with a one dimensional contour positioned exactly in 0 can be represented then by the following function



2.2.5 **Calculation of the First Derivative**

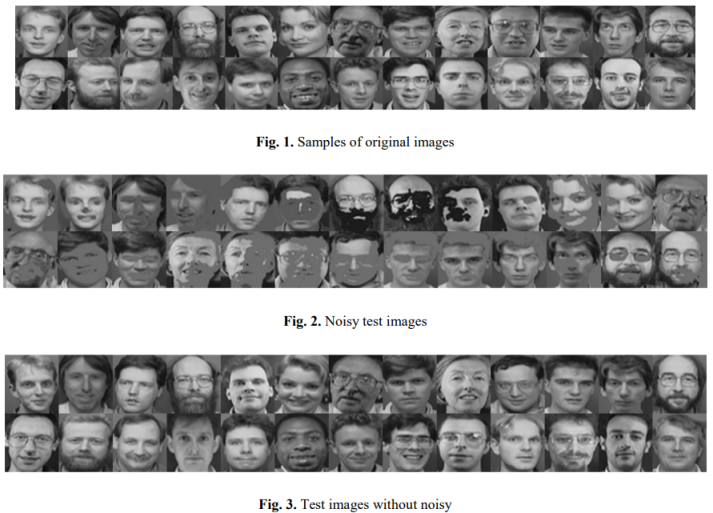
Many algorithms for the recognition of contours operate on the first order derivative of the light intensity which corresponds to the gradient of the intensity of the initial image. Based on this we search the peak values of the gradient of intensity. If I (x) represent the intensity of pixel x and I’(x) denotes the derivative (gradient intensity) to the pixel x



**2.2.2** **Calculation of the Second Derivative**

Other operator for edge detection is based on a calculation of the second order derivative of the intensity, which roughly corresponds to the rate of change of the gradient. In the ideal case-in which the intensity varies in a continuous manner-the second derivative vanishes at the points of maximum gradient. This method, however, works well only if the image is represented in a suitable scale. As explained before, a line corresponds to a double contour and then you will have a gradient of intensity on one side of the line, immediately followed by a gradient of opposite value on the opposite side. For this reason it can be expected to have large variations in the gradient images containing lines. If I (x) is the intensity value at the point x and I’’(x) is the second derivative at the point x,

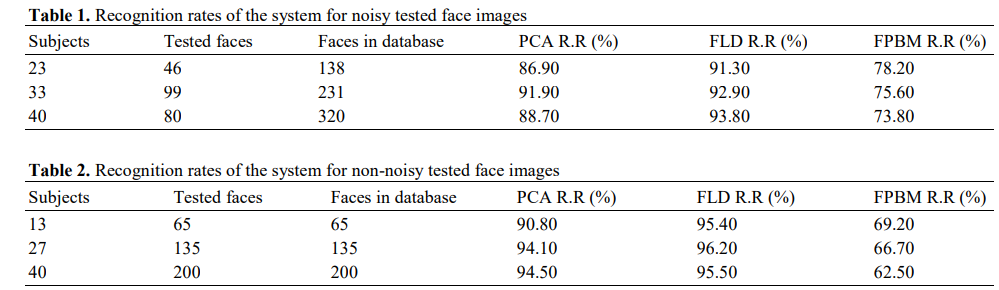




In the first experiment, 65 images of 13 persons are taken. In this image database each person has 5 different images. Number of test images that are used in this experiment is 65. Using PCA the 59 images were recognized successfully with the recognition rate 90.8%. Using FLD the 62 images were recognized successfully with the recognition rate 95.4%. Using FPBM the 45 images were recognized successfully with the recognition rate 69.2%.

In the second experiment, 135 images of 27 persons are taken. Each person has 5 different images. Using PCA the 127 images were recognized successfully with recognition accuracy 94.1%. Using FLD the 130 images were recognized successfully with recognition accuracy 96.2%. Using FPBM the 90 images were recognized successfully with recognition accuracy 66.7%.

In the third experiment, 200 images are taken. These images belong to 40 persons and each person has 5 different images. Using PCA the 189 images were recognized successfully with recognition accuracy 94.5%. Using FLD the 191 images were recognized successfully with recognition accuracy 95.5%. Using FPBM the 125 images were recognized successfully with recognition accuracy 62.5%. The simulation results demonstrate the efficiency of using of PCA and FLD methods over FPBM method in face recognition.



**2.3 FaceRecognition Approaches**

A facial recognition system is a technology potentially capable of matching a human face from a digital image or a video frame against a database of faces. Such a system is typically employed to authenticate users through ID verification services, and works by pinpointing and measuring facial features from a given image.

Facial recognition systems are employed throughout the world today by governments and private companies. Their effectiveness varies, and some systems have previously been scrapped because of their ineffectiveness. The use of facial recognition systems has also raised controversy, with claims that the systems violate citizens' privacy, commonly make incorrect identifications, encourage gender norms and racial profiling, and do not protect important biometric data. The appearance of synthetic media such as deep fakes has also raised concerns about its security.

**2.3.1 History of facial recognition technology**

Automated facial recognition was pioneered in the 1960s by Woody Bledsoe, Helen Chan Wolf, and Charles Bison, whose work focused on teaching computers to recognize human faces. Their early facial recognition project was dubbed "man-machine" because a human first needed to establish the coordinates of facial features in a photograph before they could be used by a computer for recognition.

In 1970, Takeo Kanade publicly demonstrated a face-matching system that located anatomical features such as the chin and calculated the distance ratio between facial features without human intervention. Later tests revealed that the system could not always reliably identify facial features.

Nonetheless, interest in the subject grew and in 1977 Kanade published the first detailed book on facial recognition technology.

Following the 1993 FERET face-recognition vendor test, the Department of Motor Vehicles (DMV) offices in West Virginia and New Mexico became the first DMV offices to use automated facial recognition systems to prevent people from obtaining multiple driving licenses using different names. Driver's licenses in the United States were at that point a commonly accepted form of photo identification.

**2.3.2DIFFERENTAPPROACHES OF FACE RECOGNITION**

Face recognition has attract researches in different backgrounds such as face recognition, face pattern, neural networks, computer vision, computer graphics and psychology. It is challenging method, but yet it is interesting. Some of the face recognition methods are

1. Holistic matching

2. Feature-based matching

3. Hybrid matching

1. Holistic matching methods

In this method entire face is used as a raw input to a recognition system. The holistic matching method can be classified into linear and non-linear projection methods.

Linear projection appearance-based method includes principal component analysis (PCA), independent component analysis(ICA), linear discriminate analysis(LDA) and linear regression classifier(LRC).

Non linear projection appearance- based method includes kernel principal component analysis, kernel linear discriminate analysis, and locally linear embedding. In non linear approach the input image is mapped into higher dimensional space in which the face is simplified and linear.

Hence traditional methods are applied.

In principal component analysis a number of images are taken using grey levels. Each image is mapped to a long vector of grey levels.

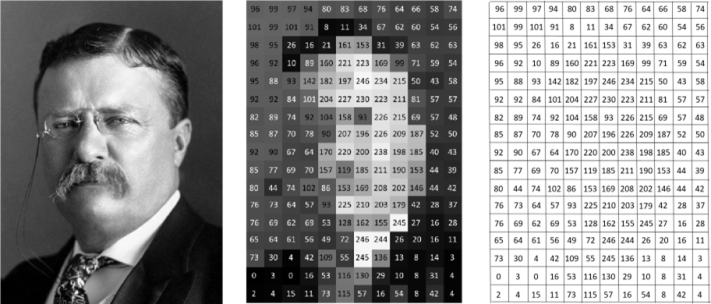
The dimensionality of each vector will be very large. The grey levels are also too sensitive to noise and lighting conditions. A possible solution to these problems are reducing the dimensionality of a space by finding principal component( Eigen vectors) to space the face and only a few significant Eigen vectors can be used to represent a face thus reducing the dimensionality.

In the first case a training set of size nxn is created. After creating training set, each image is converted into vectors. The next step involves is normalization of vector. Normalization means finding common features among various images i.e. so that each image is left behind with its unique feature. If is the normal face vector vector is the average face vector than the normalized face vector is given by

**[C= AAT];**

**{where A =1+2+3+……+n, A=N x M}**

where N are the number of rows and M is the number of columns. To recognize an image each k eigen vector is assigned weights. The distance between the input weight vector and all the weight vectors of the training set is compared. If this distance is less than threshold value than the unknown person is not matched but if it is above than the face is matched.



1. Feature based method

In this method the eyes, nose, mouth are located and features are extracted which are feed to the structural classifier. In feature based method face restoration is a big challenge. Due to large variations the system is unable to retrieve features. The extraction methods can be distinguished into

General methods based on eyes ears and nose

Feature-template based method

Structural matching methods ( geometrical Constraints on the features)

Effective features that can be used in feature based face recognition can be classified as follows:

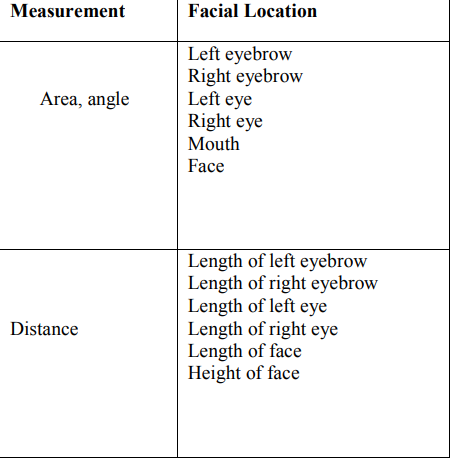


Fig :-First order feature

# Another configurable set of features that characterize the spatial relationships between the positions of the first-order features and information about the shape of the face are called second order features. [5] This approach is simple to implement and heuristics can be applied easily but it is not a robust method for face recognition. This method is heavily dependent on the external environment features.

1. Hybrid method

It is a combination of holistic method and feature based matching method. In this method 3-D images are used.

This allows the system to note the curves of eyes, nose, cheeks etc. full face can be constructed as the depth and axis of measurement gives a lot of information about the face.

This method involves detection of the face either by scanning or by photograph taken in real time. The location, angle and size of the head is positioned.

Each curve is measured and a template is made and the region outside the eye, inside the eye, region of the nose is focused.

This template is than converted into code. This code is stored in the database and than later compared with input image for recognition.

# 3 Face Recognition Techniques

# While humans can recognize faces without much effort, facial recognition is a challenging pattern recognition problem in computing. Facial recognition systems attempt to identify a human face, which is three-dimensional and changes in appearance with lighting and facial expression, based on its two-dimensional image. To accomplish this computational task, facial recognition systems perform four steps. First face detection is used to segment the face from the image background. In the second step the segmented face image is aligned to account for face pose, image size and photographic properties, such as illumination and gray scale. The purpose of the alignment process is to enable the accurate localization of facial features in the third step, the facial feature extraction. Features such as eyes, nose and mouth are pinpointed and measured in the image to represent the face. The so established feature vector of the face is then, in the fourth step, matched against a database of faces

|  |  |
| --- | --- |
| Measurment | Facial Recognition |
| Distance | Left eyebrow – Right eyebrowLeft eye – Right eyeLeft eyebrow – Right eyeRight eyebrow – Left eyeRight eyebrow – MouthLeft eyebrow - MouthLeft eye - MouthRight eye - MouthEyebrow – Side of |
| Angle | Left eyebrow - Left eye - Left eyebrowRight eyebrow - Right eye - Right eyebrowLeft eye – Mouth - Right eyeLeft eyebrow - Left eye - MouthRight eyebrow – Right eye - Mouth |

# They are catogerize in to two parts:-

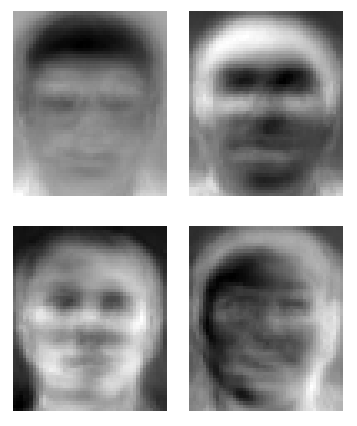
# Traditional Technique

# Deep Learning Approach

3.1 Traditional Techniques

Some face recognition algorithms identify facial features by extracting landmarks, or features, from an image of the subject's face. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with matching features.

Recognition algorithms can be divided into two main approaches: geometric, which looks at distinguishing features, or photo-metric, which is a statistical approach that distills an image into values and compares the values with templates to eliminate variances. Some classify these algorithms into two broad categories: holistic and feature-based models. The former attempts to recognize the face in its entirety while the feature-based subdivide into components such as according to features and analyze each as well as its spatial location with respect to other features.



* + 1. Eigen faces

Eigenfaces refers to an appearance-based approach to face recognition that seeks to capture the variation in a collection of face images and use this information to encode and compare images of individual faces in a holistic (as opposed to a parts-based or feature-based) manner.

The eigenfaces may be considered as a set of features which characterize the global variation among face images. Then each face image is approximated using a subset of the eigenfaces, those associated with the largest eigenvalues.

Specifically, the eigenfaces are the principal components of a distribution of faces, or equivalently, the eigenvectors of the covariance matrix of the set of face images, where an image with N pixels is considered a point (or vector) in N-dimensional space.

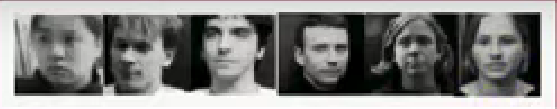
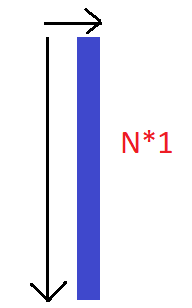


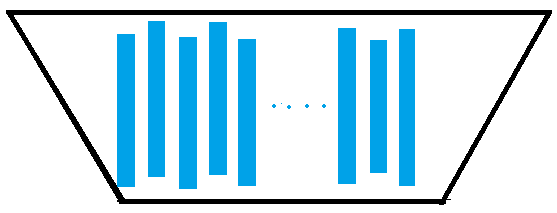
Fig: A training set consisting of M images

 N\*N image 

**Steps to find Eigen values**

Step 1:

**Convet face image in Training set to face vector**



Face vector space (



Average face()

* Faces are converted into face vector space, as shown.
* Using face vector space average face is calculated, by collecting common features.
* To get normalize face vector (subtract the mean face vector from Each face vector)
* Normalized face vector

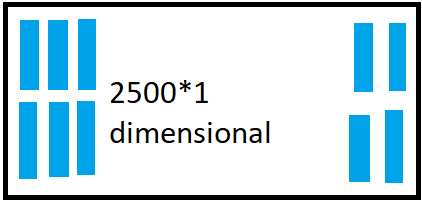
Φ*i*=Γ*i*−Ψ

* Then, the Eigen vectors are calculated, to find the Eigen vector the covariance matrix C is calculated
* It is defined as

{Where A= \*M}

C=

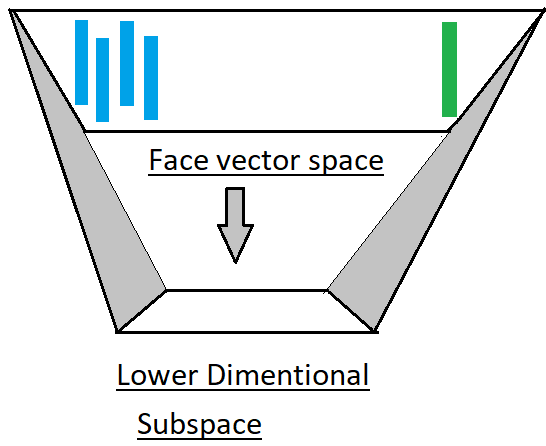
* Let’s take an example for an image whiach has a size of 50\*50.
* So, covariance matrix will be = 2500\*2500 and this will generate 2500 eigen vectors.



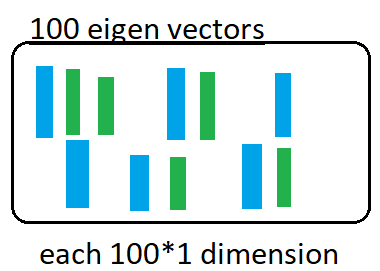
* After calculating, eigen vectors we need to find K (K is the no of eigen vector selected for training set) , where K<=M.
* For calculating the value of the dimension should be fixed(reduced).

Step 2 :

**Reducing Dimension**



* Here, a new C is introduced which is in dimentionally reduced form.
* = = M\*M = 100\*100



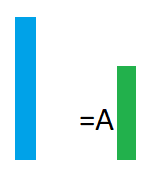
* Rest of things are calculated as previous one.

Step 3:

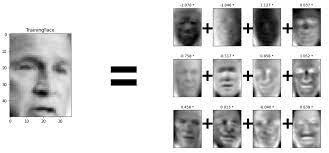
**Selecting K best eigen vectors**

* Here, K<M and can represent the whole training set.
* Selected K eingenfaces must be in the original dimension of the vector space.
* Converting lower dimensional K eigenvectors to original face dimensionally.

= =



* Blue colur denotes the higher dimension and green colour lower dimension.





Keigen faces

Selected eigen face

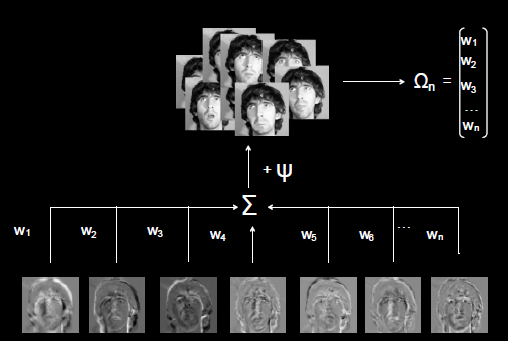


* Here, all the dimentionally reduction is done, so we will select one eigen face from the training set.

Step 4:

**Representing eigen face**

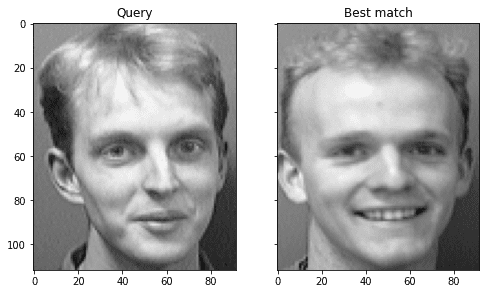
* Now each image face will be represented in linear caombination of all K eigenvectors.
* It is reprsented as Each face from training set = weighted sum of the K eigenfaces + the mean face



* A weighted vector which is the eigenfaces representation of the ith face.
* Weighted vector for each face is calculated.
* Further wighted training set is used to recognize the known faces.

Step 5:

Recognising an unknown face



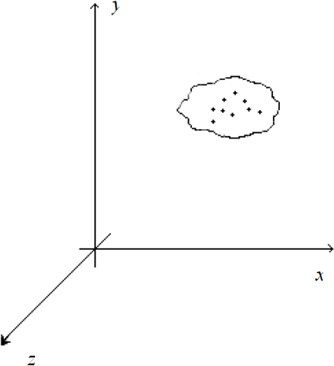
|  |
| --- |
| Normalize this face vector |

|  |
| --- |
| P Project normalize face vector onto the eigen space |

### Figure 3: image space

Weight vector of input image

* Now, we will compare the values with the training set values.
* This calculation include distance between the inputs weight vector and all the weight vectors of training set.
* So, if the distance is less than the threshold the it will recognise it as unknown face.(threshold means the most simillar features)
* And if the distance is greater than the threshold AI will recognize it as known face.

 Figure 2: Eigen-vectors or Eigen-face

Hence we can infer that the whole of image space is not optimal for face description. Hence we aim to build a face space which better describes the face and the basis vectors of this face space are known as principal components. The dimensions of the image space are equal to the total number of pixels of the image. The dimensions of the face space cannot be determined but are far less than the dimensions of the image space.

In order to perform face recognition of an unknown face, we must have a database of face images of all the students in the classroom. At least 10 (120\*120) face images of each student with variations in expressions and angle must be fed into the database. The Eigen faces algorithm works independent of the color of the skin; hence the face images are converted to grayscale and processed with histogram equalizer before being stored into the database.

The second step in the Eigen faces algorithm is to calculate the mean image of all the face images present in the database. This mean face image is then subtracted from each of the training face image.



where corresponds to the nth image in the database.



Next we calculate the eigen-vectors that correspond to the



Figure 4: reconstructed face from eigen-faces

highest eigen-values of the co-variance matrix of the set of training images.



The eigen-vectors can be thought of as a set of features which characterize the variation between face images. Eigen-vectors have ghostly face appearance; and hence are also referred to as eigen-faces as shown in Fig.3.

The technique of Principal Component Analysis (PCA) is used to generate the set of eigen-vectors (eigen-faces). The number of eigen-vectors is equal to the number of training face images in the database as shown in Fig.4.

The face images in the training set are then re-constructed by a weighted sum of a small set of characteristic images.

All the training face images are projected onto the face space. The test face images are also projected onto the face space and the distances between the test image and all the training face images are calculated. The Euclidean distance is used as the basis for calculating the distance between the face images. The training face image for which the Euclidean distance is the minimum is identified as the unknown face.

# Installing the lens system (optical zoom)

A webcam does not provide optical zooming capability; hence it does not prove to be very useful since the detected faces in the captured frames must have dimensions greater than 120\*120 pixels (standard size). A webcam is able to do so only if the object (the person) is close enough; but in a classroom the students are situated at an appreciable distance from the whiteboard/webcam and hence the webcam is not very useful.

In order to overcome this limitation a lens mechanism is installed in front of the webcam’s CCD sensor. The lens mechanism consists of a series of lenses divided into two sections- one for zooming and the other for focusing. Hence we can now zoom and focus accordingly by adjusting the zoom lens and focus lens respectively and the distance of the object from the webcam is no longer a limitation.

But with this advantage comes another problem of adjusting the zoom and focus lenses respectively separately. So in order to overcome this problem two separate servo motors are used and their armatures are connected to the lenses with the help of strings similar to a pulley system as shown in Fig.6. The servo motors are deliberately used for this purpose because they are easy to install and use, plus they are very compact.

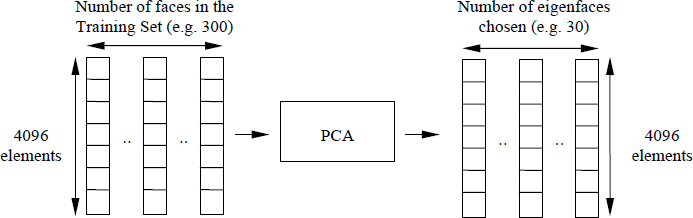


Figure 5: Principle Component Analysis (PCA)

and light. Also the servo motors can be controlled precisely and can provide feedback which helps us to know their present position.

**3.1.2FisherFace**

Fisherfaces, also known as Fisher's linear discriminant analysis (FLDA) or Fisherface algorithm, is a technique used for face recognition and classification. It is named after the mathematician and statistician Ronald Fisher, who developed the method.

Data is assumed to be uniformly distributed in each class.

The aim of fisherface is to maximize the ratio of he between class scatter matrix and with in class scatter matrix .

It produces good and error free result even when the illumination is varying.

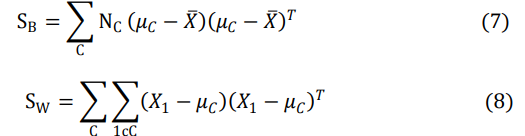
Here’s a general overview of how Fisherfaces work:

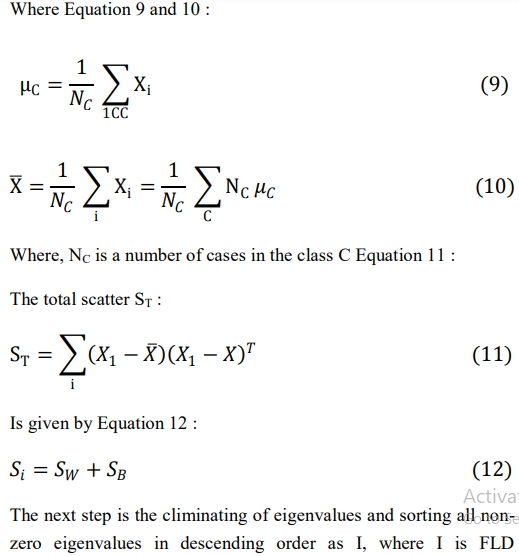
1. Dimensionality Reduction: Fisherfaces employ dimensionality reduction techniques to reduce the high-dimensional face image data to a lower-dimensional subspace. The goal is to find a set of basis vectors (eigenvectors) that represent the most discriminative features among the face images.
2. Data Preparation: The first step involves collecting a dataset of face images, where each image is associated with a unique identity or class label. These images should be preprocessed, typically by normalizing the image size, adjusting lighting conditions, and aligning the facial landmarks.
3. Classification: Once the Fisherfaces subspace is obtained, classification algorithms, such as nearest neighbor classifiers or support vector machines (SVMs), can be applied to recognize or classify faces. During classification, new face images are projected onto the Fisherfaces subspace, and the distance or similarity between the projected image and the known class prototypes is calculated for identification or classification.
4. Fisher's Linear Discriminant Analysis: Fisherfaces utilize Fisher's Linear Discriminant Analysis (FLDA) to determine the most discriminative directions in the reduced subspace. FLDA seeks to maximize the ratio of between-class scatter to within-class scatter. This means that the projected face images from the same class are clustered closely together, while face images from different classes are separated as much as possible.

Fisherfaces have been widely used in face recognition research and applications. However, it's worth noting that more recent advancements in deep learning, such as convolutional neural networks (CNNs), have achieved significant improvements in face recognition accuracy and are commonly employed in state-of-the-art face recognition systems.

Using FLD algorithm the calculations of the “within scatter matrix” and the “between scatter matrix” are performed to obtain the projected fisher images that are used in recognition (Welling, 2005; Tucker et al., 1997; Mika et al., 1999). Fisher’s linear discriminant function J is defined using covariance matrices. FLD considers maximizing the following objective Equation 6:

• SB is the "between classes scatter matrix" and • SW is the "within classes scatter matrix" They are defined as Equation 7 and 8 :

****

****

* + 1. **Local Binary Patterns (LBP)**

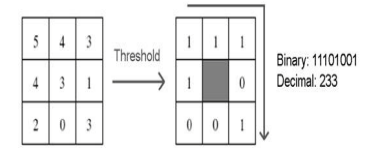
This relative new approach was introduced in 1996 by Ocala et al.. With LBP it is possible to describe the texture and shape of a digital image. This is done by dividing an image into several small regions from which the features are extracted.



These features consist of binary patterns that describe the surroundings of pixels in the regions. The obtained features from the regions are concatenated into a single feature histogram, which forms a representation of the image. Images can then be compared by measuring the similarity (distance) between their histograms. According to several studies [2, 3, 4] face recognition using the LBP method provides very good results, both in terms of speed and discrimination performance. Because of the way the texture and shape of images is described, the method seems to be quite robust against face images with different facial expressions, different lightening conditions, image rotation and aging of persons.

Principles of Local Binary Patterns

The original LBP operator was introduced by Ojala et. This operator works with the eight neighbors of a pixel, using the value of this center pixel as a threshold. If a neighbor pixel has a higher gray value than the center pixel (or the the same gray value) than a one is assigned to that pixel, else it gets a zero. The LBP code for the center pixel is then produced by concatenating the eight ones or zeros to a binary code.



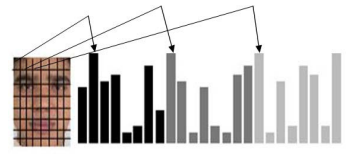
To implement the face recognition in this research work, we proposed the Local Binary patterns methodology. Local Binary Pattern works on local features that uses LBP operator which summarizes the local special structure of a face image [11]. LBP is defined as an orders set of binary comparisons of pixels intensities between the center pixels and its eight surrounding pixels. Local Binary Pattern do this comparison by applying the following formula:



Where ic corresponds to the value of the center pixel (𝑥𝑥𝑐 , 𝑦𝑦𝑐𝑐 ), in to the value of eight surrounding pixels. It is used to determine the local features in the face and also works by using basic LBP operator. Feature extracted matrix originally of size 3 x 3, the values are compared by the value of the centre pixel, then binary pattern code is produced and also LBP code is obtained by converting the binary code into decimal one.

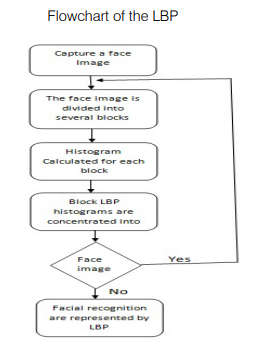
Feature Vectors

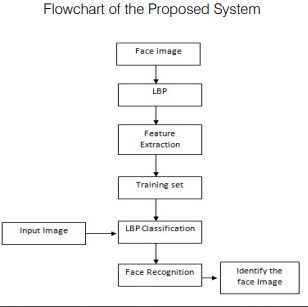
Once the Local Binary Pattern for every pixel is calculated, the feature vector of the image can be constructed [10]. For an efficient representation of the face, first the image is divided into K2 regions. In figure 1.7 a face image is divided into 82 = 64 regions. For every region a histogram with all possible labels is constructed. This means that every bin in a histogram represents a pattern and contains the number of its appearance in the region. The feature vector is then constructed by concatenating the regional histograms to one big histogram



Face image divided into 64 regions, with for every region a histogram.

Flowcharts





**Results**

This implementation is used to test the performance of the LBP-method on different kind of face images. Several parameters, like the LBP operator (P and R), non-weighted or weighted regions and the dividing of the regions, are varied to see the influence of these parameters on the performance. For this experiment we have collected lots of face images, some of them are collected from photographs taken with a Canon Power shot A610 camera and some are taken from A4Tech webcams. And also collected face images from the face database . In the proposed algorithm, different types of face images have been recognized.

Based on algorithm, the face image of an unknown identity is compared with face images of known individuals from a large database. In the figure 1.8 we can see the input facial images used for input for face recognition are given below:



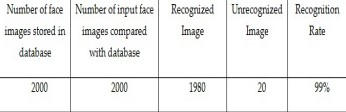
Input Facial Image

And also in the figure we can see the facial images that are stored in the database which compared with the input facial images. If the input face images are found or the more similarities face images are matched in the database then we say the face image is successfully recognized.



Facial image from Database

The following table shows overall face recognition rate:



Recognition Rate of research.

In the table the recognition rate is above 100%. We recognize the face images from the database face images by comparing between input face image and database image. From the experimental result, it is seen that the research satisfies all the requirements to recognize the face images.

1. **Challenges in Face Recognition**

Face detection and recognition are two related tasks in digital image processing that have many applications in security, biometrics, social media, and entertainment. However, despite the advances in deep learning and computer vision, these tasks are still challenging and limited by several factors. In this article, we will explore some of the current issues and obstacles that face detection and recognition research faces and how they affect the accuracy, robustness, and efficiency of the algorithms.

As the range of application is expanding day by day, the complexity of the system is increasing as well. This in fact affects the efficiency of the system. In this section of the paper we shall discuss the different challenges of face recognition systems that are present today. These challenges are related to the face image which is given as the input to the system. The algorithms used or this process varies from application to application. There are many reasons that are responsible for variation in faces. These sources of variation are classified into two main factors. They are:

1. Intrinsic factors: It is due to the physical nature of the face and not dependent on the observer. Intrinsic factors are further divided into intrapersonal and interpersonal. Intrapersonal is caused due to variation in face appearance of an individual, for example ageing, facial expression and facial paraphernalia (facial hair, cosmetics, glasses etc.)

2. Extrinsic factors: This is caused due to the variation in face appearance due to the interaction of light with the face and the observer. This will include illumination, pose, scale and imaging parameters (resolution, focus, imaging, noise etc.).

Following are the common challenges seen in face recognition system can have while detecting a face.

**5.1**  **Variation in illumination**

Variations of illuminations could reduce the efficiency of FRS. For moderate levels of lighting of the background, face detection and recognition are much difficult to perform. Variation in illumination can vary the total magnitude of light intensity being reflected back from an object. On the other hand, higher light levels could lead to over‐exposure of the face and (partially) undetectable facial patterns. There have been many algorithms such as equalization techniques that are available now to get rid of this problem to an extent. Sometimes even multiple algorithms can be used in a face recognition system to tolerate the issue of illumination. But in case of extents, it is not desirable to depend on these techniques.



Variation in illumination

5.2.2 Pose and Viewpoint Changes

Variation in pose causes significant problems in detecting a face. Pose variation can be due to change in observing angle of the observer and also due to rotation in the head position. These variations can cause a serious problem in identifying the input image. Many of the systems can tolerate small variations such as small rotations in angles. But it will be difficult when it comes to large rotational angles. The database usually consists of face images of frontal view of the faces. Since the existing FRSs are very sensitive to pose variation, pose correction is essential and could be achieved by means of efficient techniques aiming to rotate the face and/or to align it to the image's axis.

 Pose and Viewpoint changes

* + 1. **Occlusion and Partial Face Recognition**

Variation in facial appearance can also be caused due to presence of objects that such as occlusion that partially cover the face. This makes it a difficult task for the system to classify the image. Although the face is found, it may be difficult to recognize it due to some hidden facial parts, making it difficult to recognize features. This challenge can be seen in real world application where acquiring persons talking on the phone, wearing glasses, scarf, hats etc or having their faces covered with hands.



Occlusion and Partial Face Recognition

5.2.3Aging and Facial Expression Changes

Another reason for the changes in the appearance of the face could be the aging of the human face and could affect the entire process of face recognition; if the time between each image capture is large, there will be significant changes in the person. As per various study conducted by scientists, in every 10 years there will be significant changes in an individual’s face appearance. The fig. 5 shows the change in an individual’s face at different ages. It is not just the shape and lines of a face that gets modified over time; there will be changes in hairstyles as well.

Some variation in the face images can be caused due to difference in expression influenced by the individual’s state of emotion. Therefore, it is important to recognize different facial expressions for evaluating the emotional state. Human expressions consist of macro‐expressions such as, disgust, anger, happiness, fear, sadness or surprise, and other involuntary, rapid facial patterns. These facial changes can be computed with the help of dense optical flow. Cosmetics and hair styles can also be included in this challenge as changing hair style and putting make-up can also cause variation in facial expression.



Aging and Facial Expression Changes

**5.2.5 Image Resolution**

Another important issue with face recognition system is the varying quality and resolution of the images given as input. Many factors can affect the resolution of an image. The environment, the performance quality of the acquiring system and many other reasons can be mentioned as factors that are responsible for varying resolution of the image. If the resolution is good, then the recognition process will be much easier and efficient. So we can say that resolution is directly proportional to the efficiency of the face recognition system.



Image Resolution

1. Recent Advancements in Face Recognition

Facial recognition is about to be able to accurately unmask deepfakes, match families walking through the airport all at once, and outperform other biometric modalities, according to industry experts speaking at the International Face Performance Conference (IFPC) 2022.



**6.1 DeepMetricLearning**

• Deep matric learning is a subfield of machine learning that involves learning distance metrics between data points in high-dimensional spaces.

• Deep matric learning has many applications ,such as image retrieval , face recognition , and similarity learning.

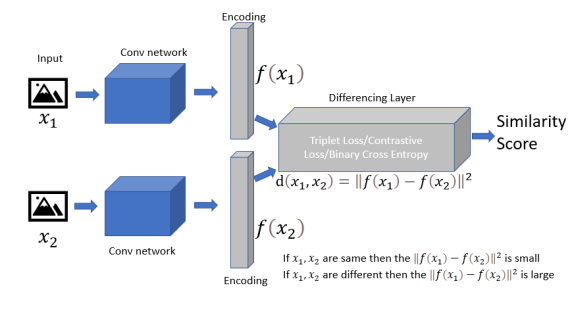
• Metric learning allows us to learn a discriminative embedding space that both maximizes inter-class distance and minimizes inter-class distance.

**6.1.1 Siamese Network**

• One popular approach in deep metric learning is the use of Siamese networks, which are neural networks that share weights across multiple identical sub networks.

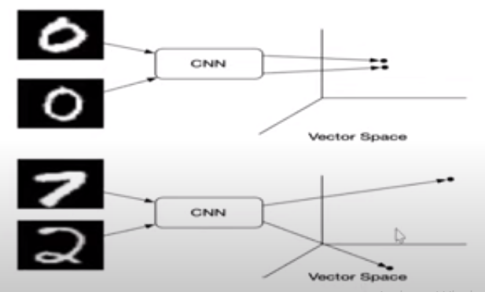
• A Siamese network consists of two or more identical subnetworks : neural networks with the same architecture, configuration and weights.

• The most commonly used loss functions are a contrastive loss function and a triplet loss function. let’s look at each of them in detail:



**6.1.2Contrastive loss**

It is (also known as pairwise ranking loss) a metric learning objective function where we learn from training data examples structured as pairs: positive pairs and negative pairs.

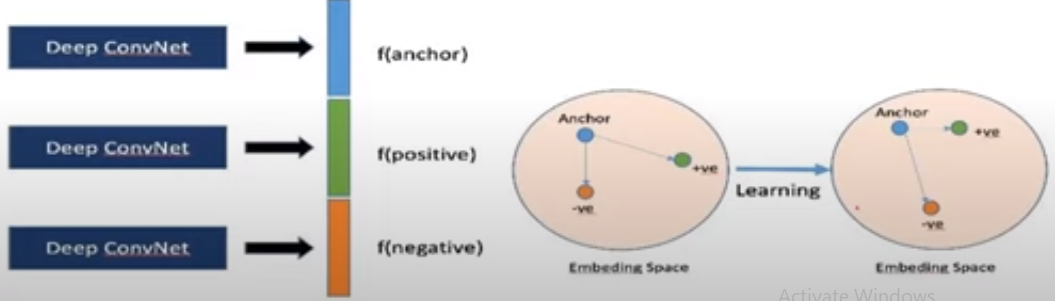


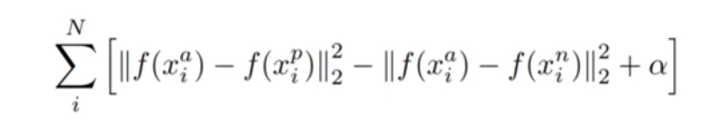
L (x1 x2 y)= Y \*D2  +(1-Y)\* max(margin – D, O)2

* the Y value is your label. it will bw 1 if the image pairs are of the same class , and it will be 0 if the image pairs are of different class.
* the D variable is the Euclidean distance between the outputs of the sister network embedding’s.
* The max function takes the largest value of 0 and the margin, m, minus the distance.

**6.1.3 Triplet Loss**

* The triplet loss function is the loss functiom that is usedto train Siamese networks in deep matric learning.
* The triplet loss function takes three input image: an anchor imagr, a positive image and a negative image.



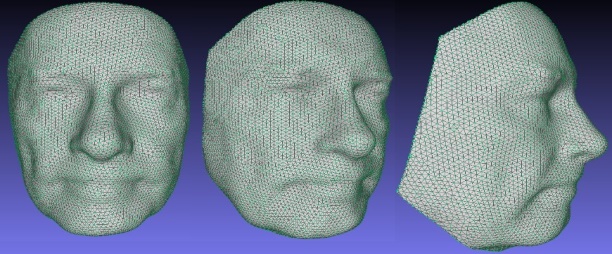


* *f(x)* takes **X** as an input and return a N-dimensional vector **w.**
* *I denotes I’th* input.
* subscript *a* indicates anchor image,  *p* indicates positive image, *n* indicates negativeimage.

**6.2 3D Face Recognition**

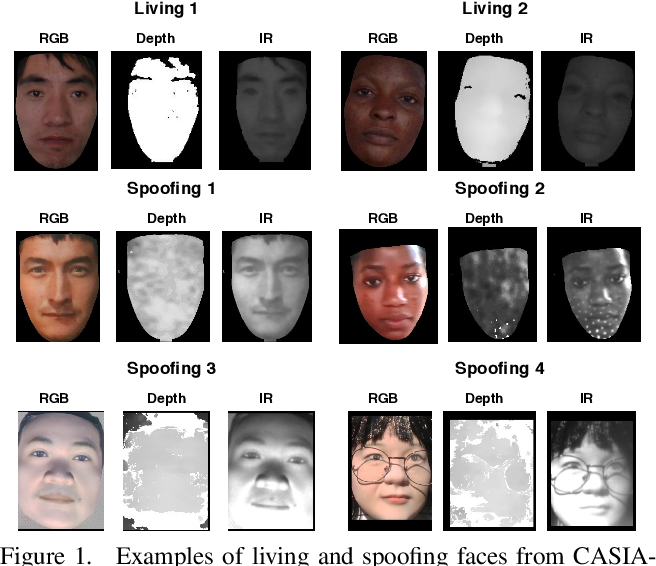
Three-dimensional face recognition (3D face recognition) is a modality of facial recognition methods in which the three-dimensional geometry of the human face is used. It has been shown that 3D face recognition methods can achieve significantly higher accuracy than their 2D counterparts, rivaling fingerprint recognition.

The main technological limitation of 3D face recognition methods is the acquisition of 3D image, which usually requires a range camera. Alternatively, multiple images from different angles from a common camera (e.g. webcam) may be used to create the 3D model with significant post-processing. (See 3D data acquisition and object reconstruction. This is also a reason why 3D face recognition methods have emerged significantly later (in the late 1980s) than 2D methods. Recently[when?] commercial solutions have implemented depth perception by projecting a grid onto the face and integrating video capture of it into a high resolution 3D model. This allows for good recognition accuracy with low cost off-the-shelf components.



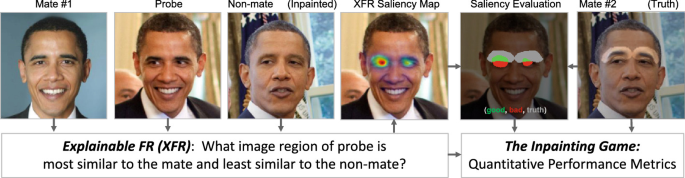
**6.3 Multi-Modal Face Recognition**

A biometric system which primarily based on the cues of unimodal biometric for individual identiﬁcation is not always meet the desired results. The concept of multimodal biometrics for human Identiﬁcation is an emerging trend. In this paper, we present state-of-the-art novel multimodal biometric system, for face recognition, which combines the similarity scores of the unimodal modalities such as appearance based and texture based techniques of face recognition, to cater the decisive results at the level of matching score. Formally, it includes the fusion of unimodal techniques to devise the multimodal models in four possible combinations such as (a) Eigenfaces and local binary pattern (LBP), (b) Fisherfaces and LBP, (c) organics’ and augmented local binary pattern (A-LBP), and (d) Fisherfaces and A-LBP. The performance of the multimodal face recognition systems is tested on the publicly available face databases such as the AT & T-ORL and the Labeled Faces in the Wild (LFW) using a new Bray Curtis dissimilarity metric. The experimental results show a signiﬁcant improvement in the performance of recognition accuracies of multimodal face recognition techniques.



**6.4 Explainable and Transparent Face Recognition**

Explainable face recognition is the problem of explaining why a facial matcher matches faces. In this paper, we provide the first comprehensive benchmark and baseline evaluation for explainable face recognition. We define a new evaluation protocol called the ``inpainting game'', which is a curated set of 3648 triplets (probe, mate, nonmate) of 95 subjects, which differ by synthetically inpainting a chosen facial characteristic like the nose, eyebrows or mouth creating an inpainted nonmate. An explainable face matcher is tasked with generating a network attention map which best explains which regions in a probe image match with a mated image, and not with an inpainted nonmate for each triplet. This provides ground truth for quantifying what image regions contribute to face matching. Furthermore, we provide a comprehensive benchmark on this dataset comparing five state of the art methods for network attention in face recognition on three facial matchers. This benchmark includes two new algorithms for network attention called subtree EBP and Density-based Input Sampling for Explanation (DISE) which outperform the state of the art by a wide margin. Finally, we show qualitative visualization of these network attention techniques on novel images, and explore how these explainable face recognition models can improve transparency and trust for facial matchers.



**7.1Face Recognition Evaluation Protocols**

E2E Protocols

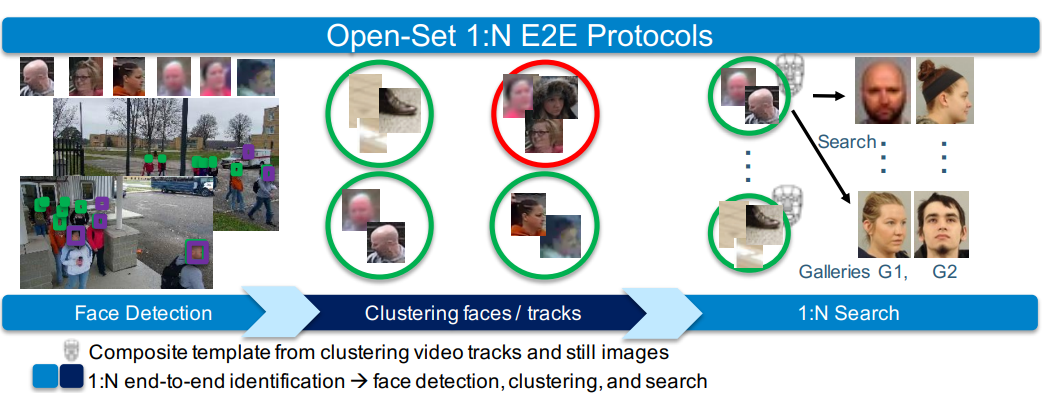
• We utilize the E2E protocols released with both IJB-C and IJB-S.

• End-to-end testing is a software testing technique that verifies the functionality and performance of an entire software application from start to finish by simulating real-world user scenarios and replicating live data.

• It is also called end-to-end testing. End-to-end testing is essential as modern software has grown to be intricate with dozens of systems interacting with each other simultaneously.

i)E2E still images/frames (IJB-C)

This is a joint detection and open-set identification protocol. Algorithms are tasked with detecting faces in still images or video frames. Detected faces are searched against curated galleries G1 and G2. This protocol resembles the operational work performed by law enforcement agencies. The protocol includes 136,734 still images and video frames.



• Joint detection and recognition, faces are detected within a pile of media including both still images and video frames. Each detected face is then searched against two disjoint galleries§§, G1 and G2. The galleries are disjoint in order to facilitate open-set identification. Joint detection, clustering, and recognition first involve face detection. Each detected face is clustered to create an identity cluster for the purpose of subject specific modeling.

ii) E2E video (IJB-C).

This protocol supports joint detection, tracking, and open-set identification. The goal is to detect faces, generate identity tracks, and create templates from the identity tracks which are then searched against curated galleries G1 and G2. The protocol includes 11,739 videos.

**iii) E2E still image and video (IJB-C).**

This is a joint detection, tracking, clustering, and open-set identification protocol and is the most challenging protocol in IJB-C. The protocol contains 31,415 still images and videos.

**iv) Surveillance-to-single/booking (IJB-S).**

Faces are detected, identity tracks are created, and templates are created from identity tracks that are then searched against the curated galleries. The galleries in the single booking protocol consist of a high resolution single mug-shot style photo for each identity. The galleries in the booking protocol consist of multiple high resolution mug-shot style photos for each identity. Both protocols include 398 videos.

**v) UAV Surveillance-to-booking (IJB-S).**

Faces are detected, identity tracks are created, and templates are created from identity tracks that are then searched against the curated galleries. The galleries in this protocol are the same as surveillance-to-booking. The probe video consists of full motion video captured from an unmanned aerial vehicle.

**7.2** **E2E Evaluation Metrics**

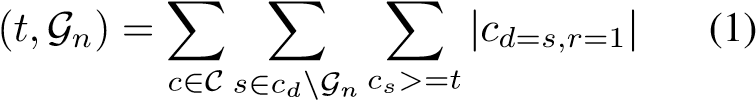
Given a set of media, M, and set of subjects, S, let all ground truth metadata consist of a media/subject mapping which yields a set of bounding boxes, represented by Nm, s, m ∈ M,s ∈ S. Let subjects of interest be enrolled in galleries Gn where Gn ∈ S. Let the total media-wise occurrences of all subjects (sightings) of interest from gallery

list, Algorithm outputs detections be defined asc ∈ C containing scores T. For this set of media, a face recognitioncd) cDsin a piece of media obtained when comparing a,m ∈ M and a candidate(cm) to a

set of detected identities ( gallery identity (cg ∈ Gn) at rank (cr).

**i) Counting False Positives**

The number of false positives is reported given a threshold, t, and gallery, Gn.

E2EFP 

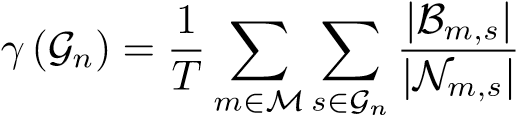
ii) False Negative Identification Rate

Given the outputs described by Figure 4, the False Negative Identification Rate (E2EFNIR) can be evaluated for all searches at all score thresholds,n t ∈ T , for a gallery con-is detaining subjects G . The minimum value of E2EFNIR termined by the faces missed by a face recognition system’s detector, B, as described below:

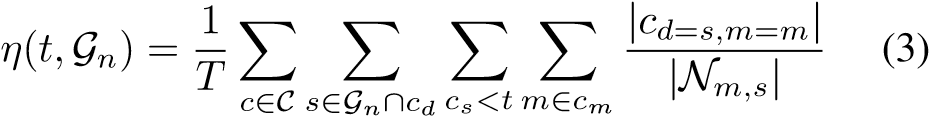
### 

### ii) False Negative Identification Rate

Given the outputs described by Figure 4, the False Negative Identification Rate (E2EFNIR) can be evaluated for all searches at all score thresholds,*n t* ∈ T , for a gallery con-is detaining subjects G . The minimum value of E2EFNIR termined by the faces missed by a face recognition system’s detector, B, as described below:

 (2)

E2EFNIR’s variable component can be defined as the sum of ratios of missed detections for an identity to the corresponding number of ground truth detections as described by Equation 3.



Finally, E2EFNIR is fully defined as follows:

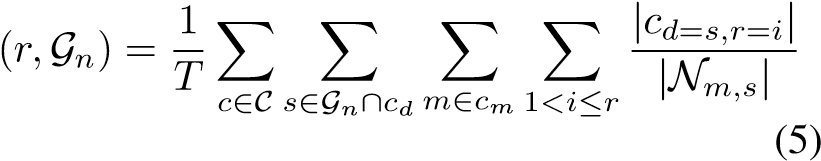
E2EFNIR (*t,*G*n*) = *γ* (G*n*)+ *η*(*t,*G*n*) (4)

### Identification Error Tradeoff Curve

E2EFNIR(*t,*G*n*) plotted against E2EFP(*t,*G*n*) yields the Identification Error Tradeoff Curve.

### Cumulative Match Characteristic

The E2E Cumulative Match Characteristic (E2ECMC) is computed for every mated candidate in each returned candidate list for a given gallery, G*n*.

E2ECMC****

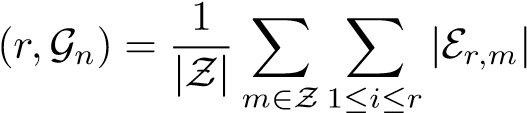
### Subject Cumulative Match Characteristic

The E2E Subject Cumulative Match Characteristic

(E2ESCMC) describes the percentage of unique, mated subjects returned on average from a collection of media for a given rank. consisting of mated subjects in Let the probe media be defined byG*n*. Let the subjects at rank*r,m*. Further,Z *r* ∈ R in media *m* ∈ Z be represented by E

to avoid counting a subject of interest multiple times (due to chimeric templates), consider only the lowest-rank

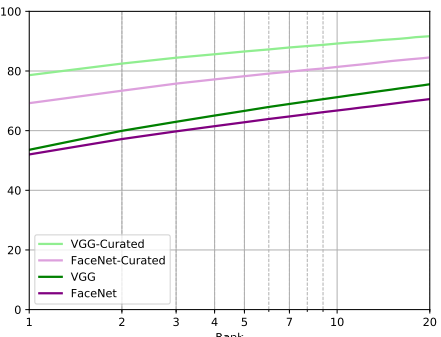
occurrence of a subject. *r, m r*+1*,m* Let *r, m*E*r, m* be governed at all*n* = ∅. Therefore, *r* E2Eby: SCMCE ∩ Eis fully defined by as follows:= ∅ and E ∩ G 6

E2ESCMC (6)

**7.3 Commonly Used Datasets**

i) IJB-C

The IJB-C E2E protocol results using VGG and Face Net are plotted along with curated results for comparison. Curated results do not require association, and make use of ground truth bounding box information and clustering for subject specific modeling. The curated results serve as a loose upper bound on performance for the E2E protocols for the specific combination of detector, clustering, and recognition algorithms.

 Fig. Average E2E closed-set performance across gallery sets G1 and G2 for the still images and frame identification protocol.

**ii) E2E Still Images and Frames**

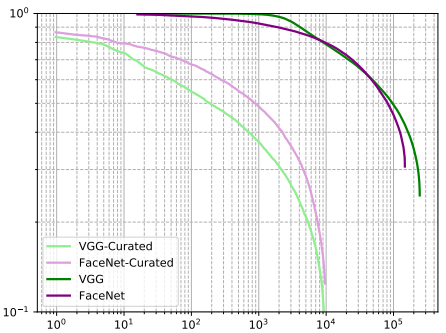
Above provides the E2E closed-set performance for both VGG and Face Net algorithms on the still images and frames protocol. As expected, curated VGG performance is roughly 30% higher at rank 1 in comparison to E2E VGG performance. Similarly, curated Face Net performance is roughly 20% higher than its non-curated counterpart. This not only highlights the difficulty associated with the E2E protocol but also illustrates that subject specific modeling can provide a significant performance gain with accurate clustering. Below figure characterizes E2E open-set identification performance on the still images and frames protocol. Curated VGG and Face Net provide a lower FNIR at a lower number of false positives in comparison to their non-curated counterparts indicating that they provide significantly better open-set performance for this protocol. Again, the curated results are provided as they serve as a rough upper bo****und on E2E performance.

Fig E2E open-set performance across gallery sets G1 and G2 for the still images and frames identification protocol.

iii) E2E Still Images and Videos

Below figure characterizes E2E closed-set performance for both VGG and Face Net algorithms on the still images and videos protocol. Both baseline algorithms perform substantially below their curated counterparts. VGG performance is roughly 5% below Face Net. We believe this difference can be attributed to detections that were discarded by our implementation of the MTCNN variant associated with VGG.

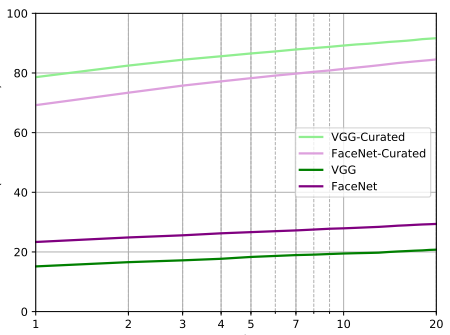


Fig Average E2E closed-set performance across gallery sets G1 and G2 for the still images and video identification protocol.

iv) IJB-S

The IJB-S E2E protocol results for Face Net and VGG are summarized. Detection and rank 1 candidate samples are shown in below figure. Overall, the UAV protocol is clearly more challenging than the other surveillance protocols as the E2E retrieval rates for both VGG and Face Net are significantly lower in comparison to their respective performance on the other protocols.

