A Comprehensive Review on Face detection using Machine learning classification

Anisha Sheikh
Research Scholar,
CSE, College of Technology and Engineering,
Udaipur, Rajasthan, India

Kalpana Jain
CSE, College of Technology and Engineering,
Udaipur, Rajasthan, India

ABSTRACT: Increased interest in biometric security systems makes face recognition a very effective research area. Now, we will take a look on basics of a face recognition system. In an image a face region is present where only facial features are present. In another terms it means we have to localize or concentrate only on the face region that means we are contemplating those parts of an image where a face may present. This research provides an up-to date survey of various existing recognition techniques but also represent precise descriptions of some methods. In addition to this other topics like issues of illumination and pose variation are also discussed in their work. The purpose of this short review paper is to present, categorize and evaluate some new face detection techniques using four conventional learning machine. The performance and the other evaluation parameters of these methods compare with each other in order to introduce significant techniques and also to state advantages and disadvantages of related works

Keywords: *Face recognition, Face detection, Machine learning, classification technique, PCA, SVM, RF*

*****

I. INTRODUCTION

A face recognition application can identify a human being in a picture or video frame digital given. Such systems are used above all in security areas together with other security technologies. Biometric authentication Facial recognition methods could be divided into "geometric" and "photometric" procedures. The First category is based on the extraction of specific characteristics of a person's image. By example, the system can analyse the relative position, dimension and shape of the eyes, mouth, nose, and the cheekbones and so on. The second category is more of a statistical approach. The base is used of image data from which the face, normalized and compressed data is extracted. Next, the image of the test is quantified in terms of that data. Among the most popular methods of recognition, I would like to mention the Eigen faces method, the Fisher face algorithm, and the Linear Discriminate analysis, the Hidden Markov Model, the Subspace of Multilinear Learning and Dynamic Link Counterpart. The latest trend in facial recognition is represented by three dimensions of recognition facial. This method uses 3D cameras to capture data about someone's face. This technology has better than classic 2D recognition, since it is not sensitive to light changes, different facial expressions, and makeup and can identify points of view, even profile. For example, the console Microsoft video games (Xbox 360) implements this new technology. It works by projecting "structured light" on the objective. Next, the system infers depth information on how the projected pattern is modified. Lately, engineers tried to create even a more powerful system by combining three cameras that point to different angles so that they could follow and recognize with high precision a person who is moving.

Robustness: This means that the biometric system should be invariant over a period of time and consequently sustain low intra-class inconsistency.

Uniqueness: This property indicates that biometric identifiers should be capable to differentiate any two persons and hence have large inter-class inconsistency

Universality: Preferably, a biometric identifier should be infatuated by every person.

Accessibility: The characteristic should be easy to acquire.

II. Applications of Face Recognition:

Face Recognition secures its own place in present times in the field of biometric security. There are several applications where face recognition is efficiently used.

i. Face Recognition is used to provide access to a smart phones, laptop, and computers. It is used for verification at an examination center or at other important places like hospitals, seminar halls etc.
ii. Now a day's in retail industry face recognition has been widely used. Retailers use this technology for Identification of customers in order to provide modified shopping experience. It can be used for target marketing scheme or for demographic section.
iii. Banking industry is the one which desperately need face recognition due to increased crimes in bank. It is used for Customer Identification, ATM access, and vault security systems in order to make banking environment highly secure.
iv. One of the most important applications of face recognition is in public surveillance. Using CCTV cameras at public places like shopping malls, cinemas etc. in order to locate missing people in the crowd especially children. CCTV

ISSN: 2454-4248
45 – 53

footage is used by the law enforcing agencies to identify criminals or terrorists.

v. In the field of Artificial Intelligence face recognition is used to identify a person and facilitates the artificial intelligent systems to carry out their work.

vi. Facial recognition can be used in time and attendance systems in various multinational companies in order to prevent the 'buddy punching' which is similar to proxy in collages and schools.

III. Face Recognition Tool and Technique

Increased interest in biometric security systems makes face recognition a very effective research area. Now, we will take a look on basics of a face recognition system. As shown in the figure (1) these are four basic steps for face recognition: First step is Localization, in this we figure out the face region. In an image a face region is present where only facial features are present. In another terms it means we have to localize or concentrate only on the face region that means we are contemplating those parts of an image where a face may present. In this localization process sometimes we may get false results due to poor quality of image which is taken as input or sometimes fault may occur due to variation in expressions. So, we have to keep these things in mind before carrying out this process.

Figure 1 example of localization of face

After localization next step is Normalization, in this we adjust the face in order to assure that all the facial features are in their exact location. Normalization process is carried out using techniques like clipping, rotation and scaling. Main requirements for normalization are constant image size, fixed eye positions, frontal orientation etc.

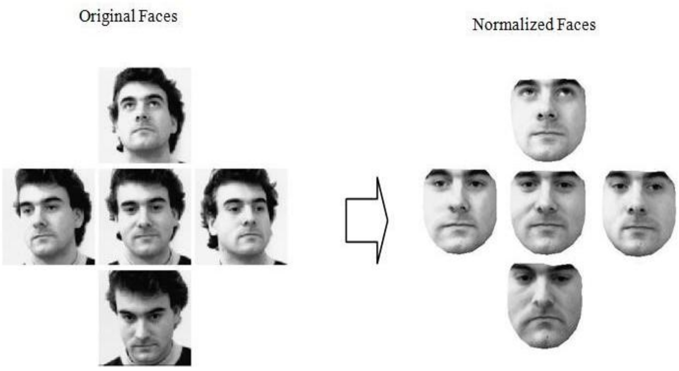


Figure 2: example showing normalization of face

Next step is considered to be as central step known as Extraction or facial feature extraction. In this extraction of various features like eyes, nose, ears, lips are done from normalized face. These extracted features are then used as input data to application. In most applications this step is considered as central step. These are several approaches used for this like geometry-based techniques, colour segmentation based techniques, template based techniques etc.

Last and final step is Verification. This is the process in which we verify the relationships between the extracted features and the database stored. Matching with database is not the only method for face recognition. There are also other techniques for finding the correlation between various parameters of facial features or by using template based model

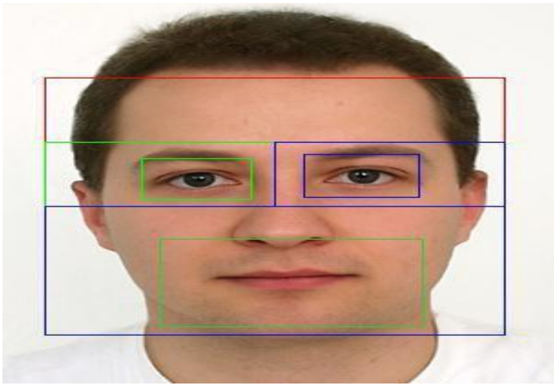


Figure 3 Figure showing extraction of facial features

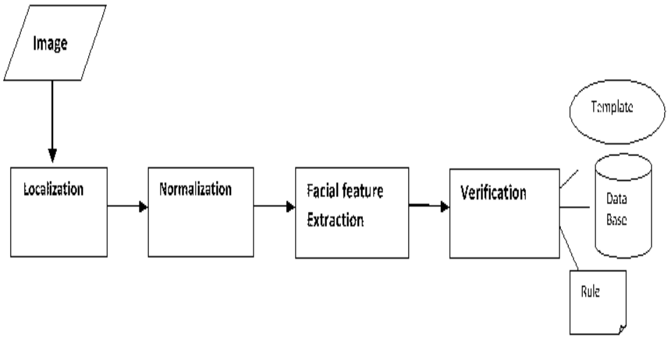


Figure 4: A layout of simple face recognition system

This biometric technology has been extensively used for recognizing impending threats like terrorist attacks. At present times it is anticipated that facial recognition will soon overtake other biometrics like fingerprint, iris recognition etc. Now, we are going to discuss about some face recognition technologies and various applications of face recognition in present world:

A. Viola-Jones Detection – Face Recognition

The Viola-Jones face detection method was proposed for the first time in a computer vision conference in 2001 by Pablo Viola and Michael Jones. His approach exceeded any existing face detector at that moment. Although you can be trained to identify many types of rigid objects, it is mostly used for face detection. Viola and Jones claim that the performance of the algorithm is comparable to previous algorithms. But, used in real time, your detector is able to run 15 times faster than the previous algorithms, without resort to techniques such as the detection of differences in the image or color of the skin. On the other hand, adding these alternative sources of information will result in achieving even more speeds high this detector is based on three concepts. The first one is known as the "Integral Image", which allows the characteristics used by this detector to be computed very fast. The second is an automatic learning algorithm, "Adaboost", which selects only the features important of the whole. The third concept is the creation of a "cascade" structure, the combination of complex classifiers, which rejects the background of the input image by passing more calculation time in the areas that may contain the object of interest. The algorithm works better in frontal objects and not so well in the lateral views, since these contribute variations in the template that characteristics cannot control well. For example, the Side view of an object must catch part of the scene change behind the profile of the object. Thus, the classifier is limited to learn the background variability. The first condition before starting to train our system is to collect the appropriate data to our situation. The "good" data is the data perfectly divided into categories. For example, do not mix inclined objects with vertical objects. The data must also be normalized the performance decreases dramatically when a classifier tries to correct the unreal variability in the data.

B. Face Recognition Using Machine Learning

As we can imagine, the main objective of computer vision is the use of machines that they emulate perfectly the human vision, being able to take measures based solely on the visual input information. However, decision making would be impossible without a technical learning. Machine Learning aims to transform the data into information. A system can learn from a set of data by extracting certain patterns and then be able to answer questions related to a new data set. When it comes to binary decisions, usually a subset of the original data is separated (from 10,000 faces of a large set, for example 9000 faces) and another for the tests (the remaining faces 1000). Using the first set the classifier builds its own model of what a face. Afterwards, the classifier will be tested in the smallest data subset to check how well it worked. If the results are bad, we might consider including some more features in the first data set or even choose a different type of classifier. TO Sometimes, jumping directly from the training to the final test could be too much. Instead of that, we could divide the samples into three: 8000 faces for learning, 1000 for validation and last 1000 for the final test. During the validation phase, we can look at the results and see their performance. Only when we are completely satisfied with this intermediate stage should we execute the classifier in the final test.

The Integral Image:- Frank Crow introduced the Summed area table for computer graphics in 1984. Later, John Lewis used this concept in computer vision. Later, in 2001, Paul Viola and Michael Jones, they used the equivalent term "integral image" within their object detection structure, to refer to a quick and efficient method to calculate the sum of pixel values in any area rectangular of a given image.

Adaboost :-The term Adaboost comes from the English "Adapting Boosting" (Adaptive Boost) and refers to a meta-machine learning algorithm created by Robert Schapire and Yoav Freund in 2003. Adaboost can be used along with other types of learning algorithms in order to improve their results. The main idea is combine the output of some weak classifiers (learners) into a weighted sum, thus creating a strong final classifier, whose error exponentially stores to zero.

Cascade filter: The Viola-Jones detector uses the Adaboost technique, but organizes the classifiers as a cascade of rejection nodes. This cascade process means that for each node, a candidate is classified as "not in class "(" not of this class ") instantly completes the computational calculation. Only the candidate who I managed to cross the entire waterfall will be classified as a face. In this way the computational cost is significantly reduces, since most areas that do not contain the object of interest are rejected in some of the stages of our waterfall.

C. Eigenfaces Algorithm – Face Recognition

The first facial recognition system was developed by Woody Bledsoe , Helen Chan Lobo and Charles Bisson in the 1960s.

They created a semi-automated program that needs an administrator to locate characteristic features of a certain image, such as the eyes, nose, mouth or ears your system calculates the relative distances between these characteristics and creates a list with the proportions specific for each subject in the database. However, this approach has proven to be quite fragile. The Eigen faces procedure was first introduced in 1987 by Kirby and Sirovich and subsequently developed by Alex Pentland and Matthew Turk, in 1991. The term "Eigen" refers to a set of own vectors, also known in linear algebra as "characteristic vectors". The main advantage of this method is that we can represent a set of images using a base formed of images "Eigen" whose dimension is much smaller than the original set. The identification can be achieved by comparing two images of both represented in the Eigen base of the set of training. The approach of Eigen Caras began with the need to find a representation of few dimensions of images of faces. Kirby and Sirovich demonstrated that the Principal Component Analysis can be Use in a group of face images to form a set of basic features. This set is known as "Eigen images" and can be used to reconstruct the original image collection. Every Original face would be reconstructed as a linear combination of the base set. We want to extract only the critical information of the test image and encode it as effectively as possible. Photographs of faces project in a space of characteristics that better illustrates the variation between the images of faces known. This characteristic of space is defined by the eigenvectors or "Eigen faces". The vector of pesos expresses the contribution of each Eigen face to the image of our entrance. These results were expanded and improved by two computer scientists, who found a efficient way to calculate the eigenvectors of a covariance matrix. Initially, an image of the face occupied a space of high dimension and the PCA method could not be applied in large sets of But Alex Pentland and Matthew Turk discovered a way to extract the eigenvectors based on in the number of input images, instead of the number of pixels. After performing the Eigen decomposition in a set of given photos, it is obtained through an analysis Statistical of the specific "ingredients" that represent our data set. The characteristics that the images of the original collection have in common they are in what is called a "middle face". By On the other hand, the differences between the images will appear in the Eigen faces. On the other hand, you can invest the process and reconstruct any initial image from the Eigen faces along with the average image. Of this mode, each face is stored as a list of values, instead of a digital photograph, saving space in the memory of the computer. The Eigen technique is also used in other types of recognition: medical images, recognition of voice, interpretation of gestures, lip reading, writing analysis. For this reason, some prefer to use the term "image Eigen" instead of "Eigen face", even though they basically refer to the same thing.

Steps to obtain the Eigen faces

1. When collecting training images, keep in mind some general rules. First, the photographs must all be taken under the same lighting conditions and then they should be normalized so that the mouths and eyes are aligned along all the images. Second, they must also have the same resolution (rxc). Each photo will be treated as a vector with rxc elements after concatenating rows of pixels. The whole set of training will be stored in a single matrix T, where each column represents an image of different entry.
2. Calculate the average image M and subtract it from each original photo in T.
3. Determine the eigenvectors and eigenvalues of the covariance matrix C (from the distribution of probability along the dimensional vector space of a face image). Each eigenvector / auto roster has the same size as the original images and can then it seen as an image as such. They are basically directions in which each input image differs from the average.
4. Select from the eigenvectors only the main components: Place the eigenvalues in reverse order and organize the eigenvectors according to these. The number of main components k is calculated by establishing a threshold value epsilon ε in the total variance.

D. **PCA (Principal Components Analysis)**

Manages to reduce the dimensionality of the problem, selecting only the coefficients responsible for maximum variation, known as Principal Components. Based on PCA, Mathew A. Turk created Eigen faces [7], a method in which the projection characterizes the facial image of an individual as the sum of the different weights adjective? Of all the factions. These factions correspond to the main components of the facial space, or what is the same, the eigenvectors of the covariance matrix associated with the highest eigenvalues.

The article [1] contains a more detailed explanation of this algorithm. PCA does not use all available information to create your transformation. Is Information is used by other methods such as FLD. This algorithm allows use information among members of the same class, looking for a transformation that maximizes the distance between classes and in turn minimizes distance between elements of the same group. The application of this procedure to facial recognition is known as Fisher faces. In relation to the Classification, algorithms of this type are considered a type of supervised learning that is, learning for which a set is available of correctly classified elements. A classifier compares a new observation with previously available examples, assigning the group (class) more like. Examples of classifiers are: - The Nearest Neighbour Algorithm (K-NN - K Nearest Neighbour). This idea appeared as a nonparametric method for the classification of patterns in a article written by Fix & Hodges [6]. K-NN matches an observation with the group of closest observations. The distance between elements can be calculate in several ways, being the most popular is the Euclidean distance. - Another example of non-parametric

ISSN: 2454-4248
45 – 53

classifier is known as Machines of Vector Support (SVM - Support Vector Machines) [2]. An SVM model consists of a representation of a group of examples as points in a space, in search of a hyperplane that separates optimally the points of a kind of another's. The hyperplane is selected so that it has the maximum separation to the examples, thus achieving a good generalization [5].

E. KNN Algorithm

The algorithm classifies each new datum in the corresponding group, as it has k neighbours closer to one group or another. That is, it calculates the distance of the new element to each one of the existing ones, and orders these distances from least to greatest to select the group to which it belongs. This group will be, therefore, the one with the greatest frequency with the smallest distances. The K-NN is a supervised learning algorithm, that is to say, that from an initial data set its objective will be to correctly classify all new instances. The typical data set of this type of algorithm consists of several descriptive attributes and a single objective attribute (also called a class).

K-NN is very sensitive to:

- The variable k, so that with values other than k we can obtain very different results too. This value is usually set after a multi-instance test process.
- The similarity metric used, since it will strongly influence the closeness relationships that will be established in the algorithm construction process. The distance metric can contain weights that will help us to calibrate the classification algorithm, converting it, in fact, into a personalized metric.

We see how the value of k influences:

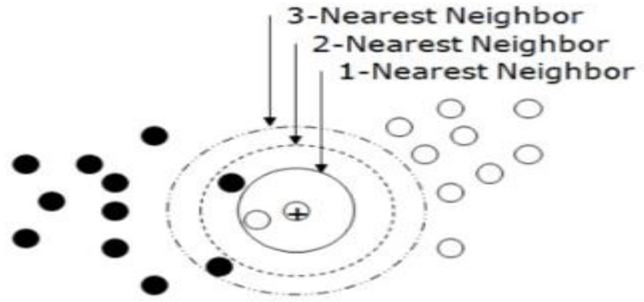


Fig 5: KNN algorithm with k influences

For k = 1 the algorithm will classify the ball with a + sign as white

- For k = 2 the algorithm has no criterion to classify the ball with sign +
- For k > = 3 the algorithm will classify the ball with a + sign as black

Its biggest weakness is the slowness in the classification process since its objective is not to obtain an optimized model, but rather each test instance is compared against the whole training data set and, it will be the goodness of the results that will determine the adjustment of aspects of the algorithm such as the own value k , the criterion of selection of instances to be part of the data set "D" of training or the same measure of similarity metric.

F. Random Forest

A random forest is a target estimator that fits a tree decision number of classifiers into several subsamples of the data set and the use of an average to improve prediction accuracy and control over assembly. Random forest reduces the variance of a large number of "complex" models with low bias. We can see the composition of the elements are not "weak" models too complex models. If you read about the algorithm, the underlying trees are planted "something" as big as "possible". The underlying trees are independent parallel models. And random selection of variables were introduced in them to make them even more independent, which makes their performance better than the ordinary bagging and right to the name of "random".

While stimulation reduces the bias of a large number of "small" models with low variance. They are "weak" as the models that are cited. The underlying elements are somehow like a "chain" or "nested" iterative model about the bias of each level. So there are independent parallel models, but each model is built based on all the ex-small weighting models. That is the so-called "impulse" one by one.

G. SVM Technique

The Vector Support Machines (SVM) are a modern and effective AI technique, which has had a formidable development in recent years, then the theoretical foundations that define these learning systems will be presented. One of the fundamental concepts in this technique is the Vector Support algorithm (VS) is a non-linear generalization of the Generalized Semblance algorithm, developed in Russia in the sixties. The development of VS brings with it the emergence of Vector Support Machines. These are learning systems that use a space of hypotheses of linear functions in a space of features of greater dimension, trained by an algorithm coming from the theory of optimization. Put more simply, the algorithm focuses on the general problem of learning to discriminate between positive and negative members of a class of given n-dimensional vectors. The MSV belong to the family of linear classifiers. Through a mathematical function called kernel, the original data is resized to look for a linear reparability of them. A characteristic of the MSV is that it maps the input vectors to determine the linearity or not of the cases which will be integrated into the Lagrange Multipliers to minimize the Empirical Risk and the Dimension of Vapnik- Chervonenkis. In general, Vector Support Machines allow finding an optimal hyperplane that separates the classes.

 ISSN: 2454-4248
45 – 53

Pseudocode to recognize faces using basic Classification technique

Steps to obtain the autorrostros (k first elements).

1. Subtract the average image M from the input image in and calculate the weight of each autorrostro Eig i.

For i = 1: k

w i = Eig iT * (In-M)

2. Collect all the previously calculated weights and form a vector W that reflects the contribution
of each autorrostro in the input image (this is equivalent to projecting the input image
on the face-space).

W = [w 1 ... w i ... .w k]

3. Calculate the distances between the test image in and each image in the database.

For j = 1: d

Dis j = || W - W j || two

4. Choose the minimum distance:

5. minDis = min j = 1: d (Dis j)

6. Determine if the input image In is "known" or not, depending on a threshold t.

If minDis <t, then Ines "known", otherwise, in is "unknown"

IV. Comparative Analysis of different Face Recognition Techniques

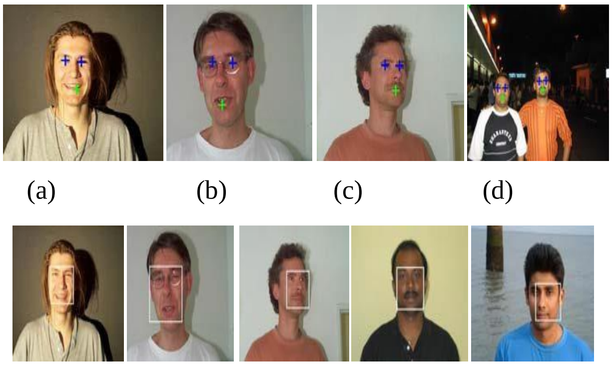
In 1999, an advanced LDA based face recognition system was presented by Li-Fen Chen, Hong yuan, and Mark liao. Linear Discriminant Analysis method [6] is used for matching of facial feature by finding the linear combination between facial features which symbolize two or more classes of objects. The linear classifier obtained from result is used generally for dimensionality reduction. LDA is nearly same as PCA. Difference between LDA and PCA is that LDA is used to describe the difference between the classes of data whereas PCA does not works on the difference instead of this the results are obtained from similarities. In this research they presented a new approach towards LDA based technique which helps in solving the small sample size problem. The orientation difference obtained after applying the proposed method was 0.0006 whereas the difference obtained through Liu's method was 75.0965. So it is obvious that the projection vector determined by their method was more stable than the Liu's method.

In a work contemplated by R. Chellappa and P.J. Phillips in 2003 outlined that this topic has recently gained momentous attention, notably during past several years. This research provides an up-to date survey of various existing recognition techniques but also represent precise descriptions of some methods. In addition to this other topics like issues of illumination and pose variation are also discussed in their work.

In 2004, work presented by Coo ray Saman and O'Connor states a hybrid technique by merging PCA with facial features like eyes and mouth are detected automatically by using RSST color segmentation [8]. In this Frontal view of image is taken in which mouth and eyes are detected. Next step is performed i.e. verification which is based on Eigen face theory to find out the relative distance between facial feature points.

A. Results produced by Different author :

Using a facial features, a hybrid solution of frontal face detection and Eigen's theories are presented in this paper. Before analyzing the PCA, it helps to deal with two requirements for this system by using the facial extraction step. Firstly, the image does not need to be searched for face-to-face per pixel position. Second, the facial detection process can usually be done in a cycle from the search space, so that the image can avoid the processing requirements on multiple scales. Using these techniques, the results below are shown in the statistics:



(a) (b) (c) (d)

Figure 6 Face detection results

In 2006, Kresimir Delac, Mislav Grgic and Sonja Grgic present their work on comparative study of PCA, LDA, and ICA on the FERET data set. In this a new methodology for correlating two CMS curves [9]. In the same year Ahonen, T. Hadid presents a unique and competent facial image presentation based on Local Binary Pattern texture features. LBP was first characterized in 1994. It has been considered as a powerful feature for texture classification. LBP divide the input image into cells, the each pixel in a cell is compared with its eight neighbours. Then the histogram is computed and normalized and then the normalized histograms are concatenated. In this approach the LBP features are extracted from divided face image and concatenated with the enhanced feature vector. Various applications and extensions are discussed in this research.

ISSN: 2454-4248
45 – 53

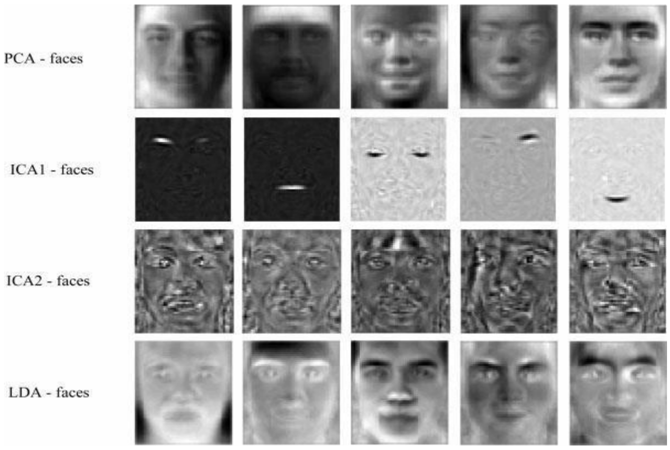


Figure 7 face representations found by PCA, ICA1, ICA2 and LDA.

In this research the three main face recognition projection methods are compared. Hypothesis testing is performed on these methods in order to calculate the differences in performance of various algorithms.

TABLE 1

Comparative study of Recognition Rate of this approach with previous approaches

	DCT+LDA + RB	DCT+LDA+ Short Distance	DCT	PCA+LDA+ Short Distance	PCA + Short Distance	Proposed Method
Recognition Rate	98.1	97.7	84.5	94.8	84	99.5

Rabia Jafri and Hamid R. Arabnia in 2009 illuminate some light on the advancement in face recognition technology by presenting a paper on survey of face recognition techniques [13]. In this paper they proposed the use of face for recognition. They present how face acts as a unique identity for an individual. Applications of this technology over other biometric technique are beautifully presented in this paper. One can refer this paper to know more about various applications and benefits of using a face for recognition. In this they both provide advantage and disadvantages of using various techniques like feature- based, statistical technique etc. This is only a review paper of various face recognition technologies.

B. **Results Obtained:**

Far-reaching experiments were carried out to personify the adequacy of the proposed approach. Fundamentally, five standard databases, i.e., the AT&T, Georgia Tech, FERET, Extended Yale B and AR have been addressed. By taking these five database experiments are carried out to justify their research work. Now, we will see the results obtained from five different databases.

C. **Results from AT and T Database**

The AT & T database is conducted at the University of Cambridge's AT & T Laboratories. This database contains 40 people, each with 10 images per person. The database includes images with different facial gestures, such as sunny or no sunny, open or closed eyes, and changes without glasses or glasses. It represents a maximum of 20 degree rotations with some size variations of around 10% size. Of these, they have accepted two evaluation protocols such as Evaluation Protocol 1 (EP1) and Evaluation Protocol 2 (EP2). The result of the results of two evaluation protocols is 2. For F1, the accuracy of 93.5% identification has been achieved in 50D properties using the LRC algorithm. Among all the best results, the feature has been reported for control and drainage (ERE) method, which is better than the 3.5% LRC system.

TABLE 2

RESULTS EVALUATED FOR EP1 AND EP2 USING THE AT&T DATABASE

	Approach	Recognition Rate
EP1	Fisher Face[24]	94.50%
	ICA[24]	85.00%
	Kernel Eigen faces[24]	94.00%
	2DPCA[24]	96.00%
	ERE[27]	97.00%
	LRC	93.50%
EP2	Fisher Face[24]	98.50%
	ICA[24]	93.80%
	Kernel Eigenfaces[24]	97.50%
	2DPCA[24]	98.00%
	ERE Sb[27]	98.30%
	ERE St [27]	99.25%
	LRC	98.75%

D. **Georgia Tech (GT) Database**

The Georgia Tech database contains 50 topics, each contains 15 images. It identifies different diversity as pose, expression, cluster background, and illumination. A 155D property structure was a sample below 15 × 15; the measurement option is illustrated in Figure 2, which reflects consistent performance on the 100D space spaces. Results obtained by the Georgia Tech Database Examination are presented at Table 3:

TABLE 3

Results evaluated for the Georgia Tech Database

Method	PCAM	PCAE	BML	DSL	NLDA
Recognition Rate	80.75	74.00 %	87.43%	90.57%	88.86 %
Method	FLDA	UFS	ERE_Sb	ERE_St	LRC
Recognition Rate	90.71 %	90.86	92.86%	93.14%	92.57 %

E. ***Extended Yale B Database and FERET Database***

Wide-ranging experiments were passed out using the Extended Yale B database and FERET Database. This Yale B database and FERET Database consists of 2,414 anterior face images of 38 persons under diverse lighting conditions. This database was alienated in five subsets; subset 1 comprises of 266 images under supposed elucidation conditions was used as the gallery, while all others were used for justification.

TABLE 4

Results obtained for the FERET Database

Experiment	Method	fa	fb	ql	qr	Overall
EP1	PCA	74.22%	73.44%	65.63%	72.66%	71.48%
	ICA I	73.44%	71.09%	65.63%	68.15%	69.57%
	LRC	**91.41%**	**94.53%**	**78.13%**	**84.38%**	**87.11%**
EP2	PCA	80.00%	78.75%	67.50%	71.75%	74.50%
	ICA I	77.50%	77.25%	68.50%	70.25%	73.37%
	LRC	**93.25%**	**93.50%**	**75.25%**	**76.00%**	**84.50%**

Subsets 2 and 3 each have 12 images, and subset 4 (14 images per person) and subset 5 (19 images of each person) illustrate the diversity of star light. All tests for the LRC system were conducted with images of 20 × 20 order samples. Results of this database are shown in Table 5.

TABLE 5

Results for the Extended Yale B Database

Approach	Subset 2	Subset 3	Subset 4	Subset 5
PCA	98.46%	80.04%	15.79%	24.38%
ICA I	98.03%	80.70%	15.98%	22.02%
LRC	**100%**	**100%**	**83.27%**	**33.61%**

3.5 AR Database

The AR database subsists of more than 4,000 colour images of 126 persons (70 men and 56 women). In this research, they deal with two fundamental challenges of face recognition, i.e., facial expression variations and adjacent occlusion.

TABLE 6

Recognition Results for Gesture Variations Using the LRC Approach

Gestures	Recognition Accuracy
Neutral	99.00%
Smile	98.50%
Anger	98.50%
Scream	99.50%
Overall	98.88%

V. Conclusion:

A system can learn from a set of data by extracting certain patterns and then be able to answer questions related to a new data set. When it comes to binary decisions, usually a subset of the original data is separated (from 10,000 faces of a large set, for example 9000 faces) and another for the tests (the remaining faces 1000). Using the machine learning classifier builds its model of what a face by different technique. Afterwards, the classifier will be tested in the smallest data subset to check how well it worked. If the results are bad, we might consider including some more features in the first data set or even choose a different type of classifier. TO Sometimes, jumping directly from the training to the final test could be too much. Lots of research have conducted their research by different classification technique but we have found that performance of SVM and Random Forest algorithm is better have other machine learning classification technique. We have seen that we can improve these classification technique by some modification in their intermediate stage. Only when we are completely satisfied with this intermediate stage should we execute the classifier in the final test.

VI. References :

[1]. J. Goldstein, L. D. Harmon, and A. B. Lesk, “Identification of Human Faces,” Proc. IEEE, May 1971, Vol. 59, No. 5, 748-760.

[2]. L. Sirovich and M. Kirby, "A Low-Dimensional Procedure for the Characterization of Human Faces," J. Optical Soc. Am. A, 1987, Vol. 4, No.3, 519-524.

[3]. M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces," Proc. IEEE, 1991, 586-591.

[4]. R. Chellappa, C. L. Wilson, and S. Sirohey, 1995. Human and Machine Recognition of Faces: A Survey, Proc.of the IEEE, vol.83, no.5, and pp.705-740

[5]. Robert J. Baron, 1981. Mechanisms of Human Facial Recognition, International Journal of Man-Machine Studies, vol.15, no.2, pp.137-178.

[6]. R. Brunelli and T. Poggio, 1993. Face Recognition: Features versus Templates, IEEE Tran. On Pattern Analysis and Machine Intelligence, vol.15, no.10, pp.1042-1052.

[7]. E. Osuna, R. Freund, and F. Girosi, 1997. Training Support Vector Machines: An Application to Face Detection, In IEEE Conference on Computer Vision and Pattern Recognition, pp.193-199.

[8]. Vladimir N. Vapnik, 1995. The Nature of Statistical Learning Theory", Springer Verlog, Heidelberg,DE.
[9]. L. Sirovich and M. Kirby, 1987. Low-dimensional Procedure for the Characterization of Human Faces, Journal of Optical Society of America, vol.4, no.3, pp.519-524.
[10]. Matthew Turk and Alex Paul Pentland, 1991. Eigenfaces for Recognition, Journal of Cognitive Neuroscience, vol.3, no.1, pp.71-86.
[11]. Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman, 1997. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Tran. On Pattern Analysis and Machine Informatics Engineering, an International Journal (IEIJ), Vol.1, No.1, December 2013 40 Intelligence, vol.19, no.7, pp.711-720.
[12]. Bernhard Scholkopf, Alex J. Smola, and Andre Bernhardt, 1998. Non-linear Component Analysis as a Kernel Eigenvalue Problem, Neural Computation, vol.10, no.5, pp.1299-1319.
[13]. M. H. Yang, 2002. Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using Kernel Methods, In IEEE International Conference on Face and Gesture Recognition, pp.215-220, Washington.
[14]. A. Jonathan Howell and Hilary Buxton, 1995. Invariance in Radial Basis Function Neural Networks in Human Face Classification, Neural Processing Letters, vol.2, no.3, pp.26- 30.
[15]. Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, and Andrew D. Back, 1998. Face Recognition: A Convectional Neural Network Approach, IEEE Trans. on Neural Networks, vol.8, no.1, and pp.98-113.
[16]. T. Poggio and K. K. Sung, 1994. Example-based Learning for View-based Human Face Detection, ARPA Image Understanding Workshop.
[17]. M. J. Er, S. Wu, and J. Lu, 1999. Face Recognition using Radial Basis Function (RBF) Neural Networks, In 38th Conference on Decision & Control,Phoenix, Arizona USA, pp.2162-2167.
[18]. C. E. Thomaz, R. Q. Feitosa, and A. Veiga, 1998. Design of Radial Basis Function Network as Classifier in face Recognition using Eigenfaces, In Vth Brazilian Symposium on Neural Networks, pp.118-123.
[19]. Y. Yoshitomi, T. Miyaura, S. Tomito, and S. Kimura, 1997. Face Identification using Thermal Image Processing, In IEEE InternationalWorkshop on Robotand Human Communication, pp.374-379.
[20]. Z. Liposcak and S. Loncaric, 1999. Face Recognition from Profiles using Morphological Operations,In International Workshop on Recognition, Analysis, and Tracking of faces and Gestures in Real-Time Systems, pp.47- 52.
[21]. Andreas Lanitis, Christopher J. Taylor, and Timothy Francis Cootes, 1997. Automatic Interpretation and Coding of Face Images using Flexible Models, IEEE Tran. On Pattern Analysis and Machine Intelligence, vol.19, no.7, pp.743-756.
[22]. Alan L. Yuille, 1991. Deformable Templates for Face Recognition, Journal of Cognitive Neuroscience, vol.3, no.1, pp.59-70.
[23]. P. Penev and J. Atick, 1996. Local Feature Analysis: A General Statistical Theory for Object Representation, Network: Computation in Neural Systems, vol.7, pp.477-500.
[24]. B. Yegnanarayana, 1999. Artificial Neural Networks, Prentice-Hall of India, New Delhi.
[25]. Simon Haykin, 1999. Neural networks: A Comprehensive Foundation, Prentice-Hall International, New Jersey.
[26]. C. M. Bishop, 1995. Neural Networks for Pattern Recognition, Oxford University Press Inc., New York.
[27]. R. J. Mammone, 1993. Artificial Neural Networks for Speech and Vision, Chapman and Hall, Cambridge.
[28]. T. Kohonen, 1988. Self-Organization and Associative Memory, SpringerVerlag, Newyork.
[29]. T. J. Stonham, 1984. Practical Face Recognition and Verification with WISARD, In Aspects of Face Processing, pp.426-441.
[30]. D. Demers and G. W. Cottrell, 1993. Non-linear Dimensionality Reduction, In Advances in Neural Information Processing Systems, pp.580-587.