Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Machine Learning Approach for Face Identification using Dimensionality Reduction Algorithm

Pragya baluni[1*], Devendra Singh[2], Bhumika Gupta[3]
[1]*Research Scholar, IFTM University, Moradabad, India.*
[2]*Associate professor, IFTM University, Moradabad India.*
[3]*Associate professor, GBPIET, Pauri Garhwal.*
[1] *(Email id: pragya.uniyal115@gmail.com)*

ARTICLE INFO

Received: 19 Dec 2024

Revised: 29 Jan 2025

Accepted: 12 Feb 2025

ABSTRACT

Introduction: Identifying faces mainly for surveillance purpose is a field of trending. Identifying faces from the image and performing recognition based on existing dataset is a challenging task. Different pose, age-based images, Blurriness, illumination are the factors which creates a challenge to perform face recognition and identification with utmost accuracy. Facial features are used to identify individual faces. Such applications if built can be used in various domains mainly including surveillance system. In past 10 years regressive research is being done in the field of computer vision to attain greater accuracy along with efficiency with less computation time. Principal Component Analysis is a statistical approach which transforms the dataset from high dimensionality to lower dimensionality keeping intact the necessary information. Eigen face is one of the most classical and hence one of the best approaches for feature extraction. Integrating PCA with CNN and SVM can lead to great results and less computation time. The study captures the details of the dimensionality reduction algorithms. A systematic comparison in presented in the paper along with the proposed model.

Keywords: Surveillance, PCA, SVM, Dimensionality Reduction, CNN.

INTRODUCTION

Identification of human faces from the dataset has become the most trending research interest which includes challenges as well. The faces with occlusion or different poses are creating difficulties in achieving accurate results. Biometrics are using face recognition applications in various fields like health industry, banking sector, legal surveillance system and many more. The face recognition system follows two major steps. First is Face Verification followed by face identification. The facial features are selected and matched with dataset and database. Figure 1 represents the various steps during face recognition process [1]. Firstly, an input image is given to which preprocessing is done. Preprocessing removes the noisy data. Post preprocessing, feature extraction in done. Various algorithms are defined for feature extraction like PCA, LDA and more. Once feature extraction is done classification is performed using classifiers for various recognition methods. SVM uses kernel classifier to perform classification on the images. At the end the output is given in form of matched image poor image identified along with name depending upon the type of application [2]. Broadly the application for recognition of faces is categorized into 2 into major categories. First appearance based and second is feature based. The detection of face is identifying human face in image dataset. The feature extraction step generates feature vectors which identifies features like shape of face, mouth, eyes, nose etc. the pixels are identified for matching faces. Finally, the faced recognition system performs comparison on extracted features with the images of selected dataset and finds the exact match. Added to this automatic face naming can be performed on identified and matched faces for better and robust face recognition system. Figure 1 below depicts the process flow for faced recognition.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

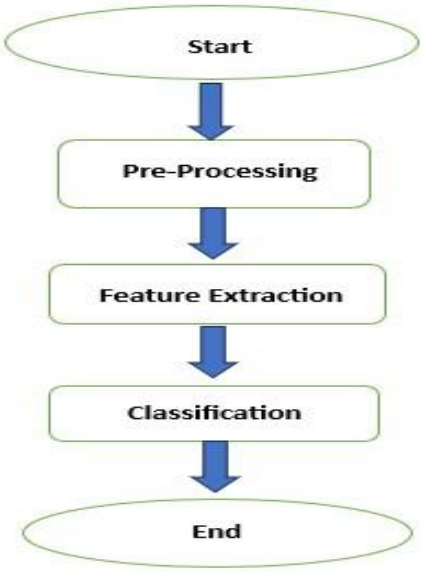

Figure 1: Face Recognition Steps

Face recognition is divided into three steps:

1. Detection of face – Firstly the face is detected in the input data which is an image the algorithm must locate the face if any in the image.
   The challenging part is detection of a moving face. The algorithm can be developed to get consistent output from the images. The points can be identified like ear, chin etc and training of images can be done to get the face detected.
2. Feature Extraction – It is the important part of any application and model for recognition of faces. This is the point where CNN plays vital role.it extracts feature and performs matching with the images in the dataset.
3. Face Recognition – The last step to be performed is the matching of features extracted from dataset. The different similarity matrix algorithm can be used to generate similarity matrix image. On the basis of which it is decided whether the face is matched or not matched with the image in dataset.
   The contribution in this study paper is to define the following:
   1. To present in-depth technologies for face recognition and identification
   2. To highlight the techniques with the respective pros and cons.
   3. To develop a model with optimized technology to enhance accuracy, efficiency and computation time.

The figure below depicts how the process of face recognition and identification works. It majorly explains the role of feature extraction. Feature extraction plays a vital role in the process of face recognition and identification.

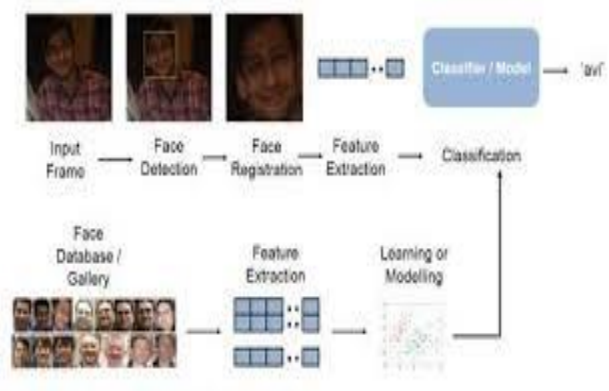

Figure 2: Process of face identification

The approaches for face identification as depicted in Figure 2, applications are broadly categorized as hoslistic methods, appearance-based, deep learning methods, Principal component analysis, LDA and many more [5]. Face identification means following two important aspects. The first is authentication and second is recognition. Both steps play a vital role in whole process of face recognition.

Deep leaning convolutional neural network is one of the most widely used machine learning tool for face recognition. CNN can extract features from the image followed by classification of the images. CNN can perform on complex dataset also with less computation time [10].

METHODS

Dimensionality Reduction Techniques

In a study by Sharma and Pattereh [4], face recognition was conducted using PCA. The approach also used ANFIS system for recognition of faces along with PCA approach. A good accuracy rate was observed when the application was executed. Principal Component Analysis (PCA) is a widely used machine learning technique for exploratory data analysis and predictive model creation.

Hamid Reza Yazdani, Ali Reza Shojaeifard [6] mentioned the eigenface and PCA for face recognition. As per the study increasing number of eigen vectors increases the chances of the errors in the process. Due to this many faces were incorrectly identified.

The integrated approach using PCA and eigen vector was accurate to around 85% whereas error was increasing with bigger dataset. The major drawback of the paper was poor identification in case of pose invariant.

Mostly pose estimated is done by PCA which reduces the dimensions of the data without losing the important features. PCA is used for Pose estimation of different variations in faces [9]. PCA is used to build projection vectors, which are then subjected to face recognition methods. By estimating the likelihood between the mean faces and the probe faces, the pose orientation of the probe picture can be determined. The mean face with the highest probability is likely to have the same pose orientation as the probing face [7,8].

It is clear that the eigen face and eigen vector approach, based on PCA, is among the most straightforward and successful feature extraction methods. Images can be converted into set of small dimensions of facial features using this method, which includes the primary components of the initial training set [11]. The mean image is determined as

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i, \quad i = 1, 2, \ldots, m.$$

The co variance matrix is as follows:

$$C = \frac{1}{n} QQ',$$

As per the finding [11], face images during identification process faces factors like noise which lead to missing important information in the matrix formed. This is the result caused due obstructions in front of the image or face [12]. Hence improving the image quality is the only alternative to improving the quality of the system. There are numerous machine learning and dimensionality reduction techniques for the same. The images need to be reconstructed with greater accuracy and better optimization approach. Although as per the study [13], statistical methods have proven good accuracy rate in such scenarios.

The final stage involves classification step. A classifier is used to compare the training set with the test.

As the study suggested [13], PCA is technique designed and executed for analysis of interrelated dependent variables. The aim is to achieve useful information and patterns from the dataset.

Eigen Face

Assume for the moment that a dataset of n pictures with identical pixel values. Every image is vectorised, and a matrix is created, with each matrix element—for example, matrix A—representing a greyscale value pixel. We now know that every matrix may be subjected to PCA, which yields specific vectors. The method aids in attaining high precision.

Changing the weight vector results in different face images since the dataset's eigenface matrix remains constant. As a result, although the photos of the same individual are not identical, their weight vectors are comparable. Consequently, the distance between two weight vectors (such the L2-norm) can be used as a measure of how similar two images are [14].

Algorithm

1. Every single image in training set is a linear combination of eigen image that are base for the dataset to be trained.
2. Weights are used to show case the image in smaller sub space (concept of dimensionality reduction).
3. For the purpose of identification, the estimated weights are done on test image and compared to that with the weight of training dataset.

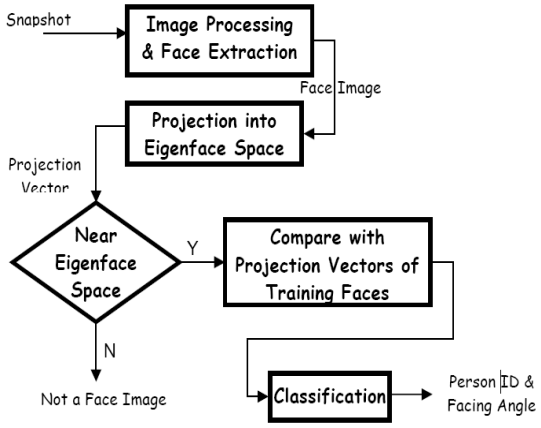

Fig 3: Flowchart of eigenface

The flowchart as mentioned in Figure 3, explains the steps of how the eigen face works for face recognition approach. In this approach by taking larger eigen values of numerous pictures from the dataset, eigen face are calculated. Then the new images are calculated using weights based on type of image given as input. This approach identifies and extracts characteristic features from the face[16].

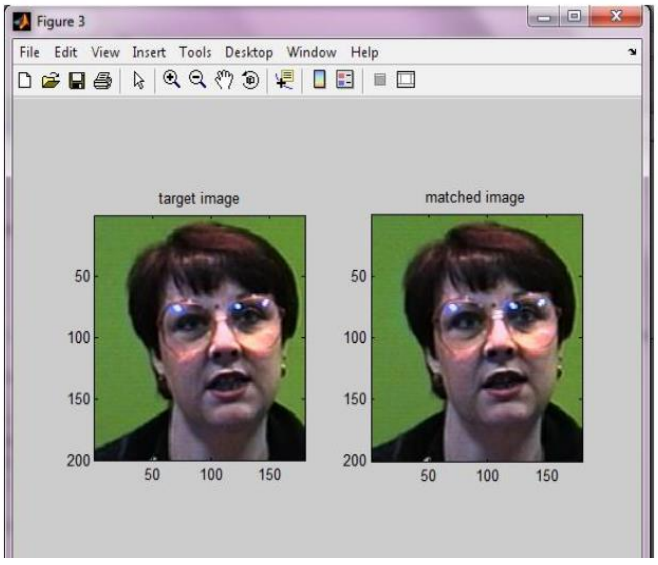

Figure 4: Matlab Implementation

Principal Component Analysis

PCA or principal component analysis, is the foundation of the eigenfaces approach. In a study [18], a group of original face photos were considered and the optimal vector scheme for compressing photographs was determined In order

to solve the face identification and recognition problem, a projection technique was used for a subspace in pattern recognition. One goal of PCA is to substitute uncorrelated vectors of smaller dimensions with correlated vectors of larger dimensions. The computation of a basis for the data set is an additional goal. The primary benefits of the PCA include its low noise sensitivity. The Eigenfaces method's approach is to extract the face's distinguishing traits and then display the face as a linear combination of the so-called "eigenfaces" that were discovered during the feature extraction procedure. Eigenfaces is a popular face recognition technique that extracts the most important traits from a collection of facial photos using PCA. The technique lowers the dimensionality of the pictures by converting them into a space with fewer dimensions while keeping the key characteristics. The most important patterns in the photographs are represented by the covariance matrix of the image collection, or eigenfaces. The technique represents the faces in the training set using eigenfaces during the identification step and then compares these representations with a fresh input image. To get the closest match, the algorithm measures the separation between the eigenfaces of the training set and the input picture [19]. Figure 4 displays how face is identified and recognised.

Pattern recognition makes extensive use of the Principal Component Analysis, a projection technique to a subspace. One of PCA's goals is to swap out large-dimensional correlated vectors for smaller-dimensional uncorrelated vectors. A face recognition system's objective is to distinguish input signals (image data) into many classes (people). This is crucial for a number of issues, including criminal identification, human-computer interaction, picture and video processing, and more. Each input signal has some patterns. Common objects like eyes, mouths, and noses, as well as the relative distances between them, are visible in all input photos.In the field of facial recognition, these shared characteristics are referred to as eigenfaces [29] (or primary components in general).

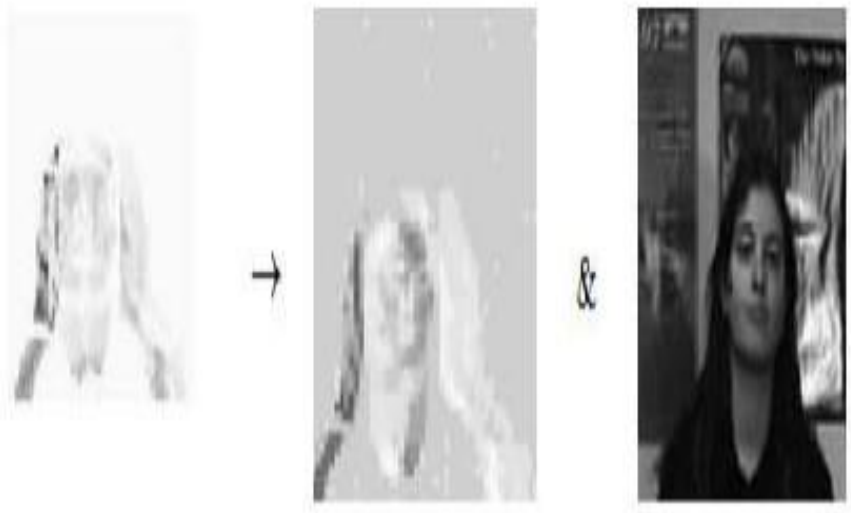

Figure 5: Eigen face representation

The above figure 5 shows how the eigenface algorithm works.The above figure displays the mean image and eigen face of the person selected from the dataset. Co variance matrix with the highest eigen value is calculated.

Variation in applications and algorithms helps in research to be redone so that it can be recorded and shared with other sources with the aim of conformation of accuracy and authenticity of results[30].

As per the study[31], following observations were made:

Table 1
Recognition rates of FaceNet algorithm subject to occlusions and expressions constraints across the adopted classifiers.

Occlusion rate	SVM	EUC	CB
30%	31.48%	36.58%	36.57%
40%	21.67%	25.0%	25.0%

Table 2
Recognition rates of FaceNet algorithm on expression variant face images after de-occlusion.

Occlusion rate	De-Occlusion method	SVM	EUC	CB
30%	MICE	63.42%	81.48%	80.84%
	MissForest	64.81%	83.33%	83.80%
	RegEM	63.43%	81.48%	81.94%
40%	MICE	58.79%	73.61%	74.08%
	MissForest	62.80%	79.52%	78.59%
	RegEM	51.85%	67.60%	67.60%

It has been observed that used face net algorithm is comparatively complex to compute results. Hence PCA based approach with SVM classifiers recommended which has lower computation complexity. Table 1 and Table 2 depicts how various mechanisms and techniques helps in face recognition.

Support Vector Machine

Three types of machine learning algorithms can be distinguished: supervised, unsupervised, and reinforcement learning. Among the best classification techniques found in the literature are artificial neural networks, the Support Vector Machine (SVM) algorithm, decision trees, the Bayesian classifier, and the k-nearest neighbours' classifier. One of the most well-known techniques for maximising the anticipated outcome from these procedures is SVM[32].SVM is a supervised learning method that may be applied to outcome prediction and classification. Finding a single hyperplane with the largest margin that can divide the classes linearly is the aim of SVM. When there is a limited amount of training data available, it is impossible to ensure the optimal response when employing a large number of statistical tests. The eigenface method is predicated on the principal component analysis, is among the most effective template matching techniques. (PCA) for facial recognition and representation.

Some of the most prominent studies and observations are mentioned below for the better understanding of the Techniques for dimensionality reduction approach.

The below table 3 mentions various techniques and respective achievements based on the past studies being carried out.

Table 3: Techniques and achievement rate

S No.	Techniques	Achievement rate	Reference
1	PCA	79.65%	[21],[8]
2	PCA+RCA	92.34%	[22]
3	Independent Component Analysis	69.40%	[19]
4	Gabor +Modular 2DPCA	96.5%	[23]
5	Gabor + 2DPCA	94.5%	[25]
6	Modular 2DPCA	92.5%	[23]
7	SVM	85-92.1%	[18]
8	Neural Networks	93.7%	[26]
9	Eigen Faces Method	92-100%	[27]
10	PCA, SVM	96%	[12]
11	Eigenfaces, PCA	98%	[6]
12	PCA, LDA, KNN,	97%	[28]
13	Machine-learning classifiers with Principal Component Analysis (PCA)	92%	[32]
14	Principal component analysis (PCA), eigen vector of PCA and Convolutional Neural Network (CNN) along with eigen vector	89%	[33]

PROPOSED MODEL

In the study it is observed that various techniques and approaches are currently being deployed to make a robust systema and application for face reognition and identification system with 100% accuracy along with less complexity and less computation. Also, a general observation is that, yet a system is to be developed with gives most accurate results with the dataset having imahes with different color, pose or occlusion. Pose is creating a big challenge in achieving a system with 100 % accuracy. The paper represents smajorly dimensionality reduction approaches. It includes the execution of PCA on different datasets along with eigen face and eigen vectoy, Support vector machine, neural network etc. Out of all the mentioned techniques eigen face and PCSA have given the results better as compared to other techniques in case of dimensionality reduction. But contrary to this the results are comprehended in case of occlusion and pose. For this matter the suggested approach is to combine PCA along with SVM so that the challenge of occlusion and pose can be eliminated.

The proposed algorithm is as follows:

Input: Set of images
Set of feature vectors
// feature vectors for angle as gamma, magnitude as theta, histogram as
lambdaOutput: Identify the persons and name them
Process
Step 1: Pre-processing

- Initiate the image loading and convert them from RGB to greyscale
- Resize the images to x*y dimensions

Step 2: Training of images and set them as matrix X
Step 3: Initialize gamma, theta and lambda vectors to get values of respective
featuresStep 4: For each matrix value compute the angle, magnitude and
histogram value
end for
Step5: For each cell in matrix compute angle, magnitude and histogram
valueend for
Step 6: Set threshold = [0, 180]
If threshold >10
&& threshold<180Step 7:
Obtain histogram channel
End if
Step 8: Compute Normalize value= Normalize value for each
matrixStep 9: Assign calculated feature vectors
Step 10: Normalize the columns of P to have unit t^2norm
Step11: When norm factor is resolved generate the test track of images of particular person whose images are
trained in step 2.
Step 12: When a test track is given identify the person, recognize them and name them.

With this proposed approach results with less complexity and less computation time can be achieved with better accuracy rate.

RESULTS AND DISCUSSION

The study is sought to understand the dimensionality reduction approach and techniques. Although there are numerous techniques execute face recognition approach but identification of face due to occlusion and pose variation is a challenge. Most of the times the images are not recognised as major part of the information is lost due data preprocessing. Hence dimensionality reduction approach like PCA along with SVM concludes with greater accuracy rate for face recognition

and identification.

Below is the sample image dataset considered for the study. It includes same face at different angles and positions.

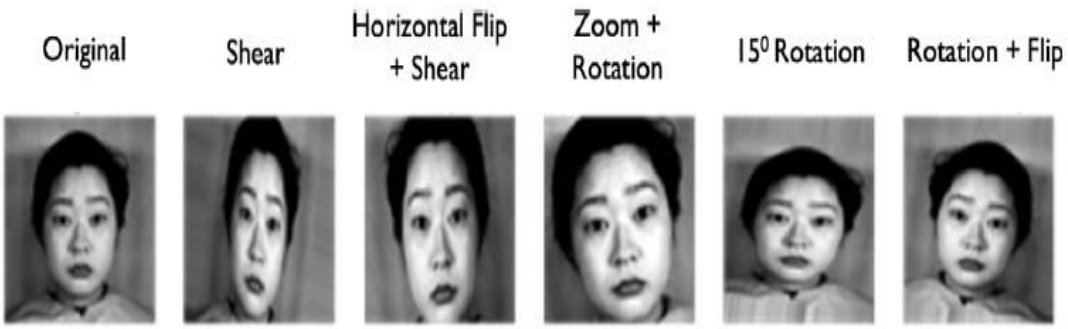

Figure 6: Sample image dataset

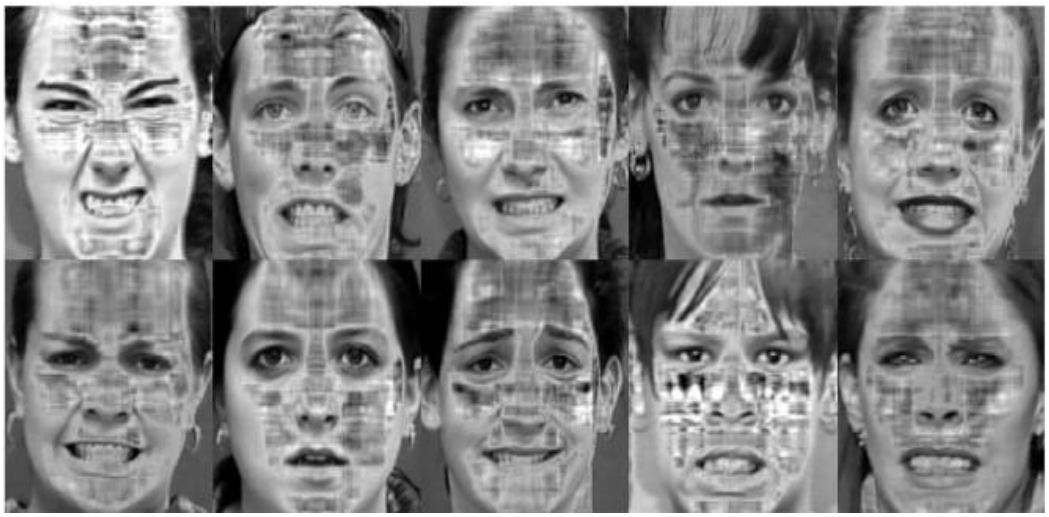

Figure7: Sample dataset for reconstructed images.

The above figure 6 and Figure 7 is the example for sample images taken into consideration for face recognition process. The graph mentioned as Figure 8 below explains that PCA if integrated with SVM or eigen face approach is showing good results with parameters like accuracy, computation time and complexity, but is not completely resolving the issue of occlusion and pose. Figure 9 displays the GUI representation.

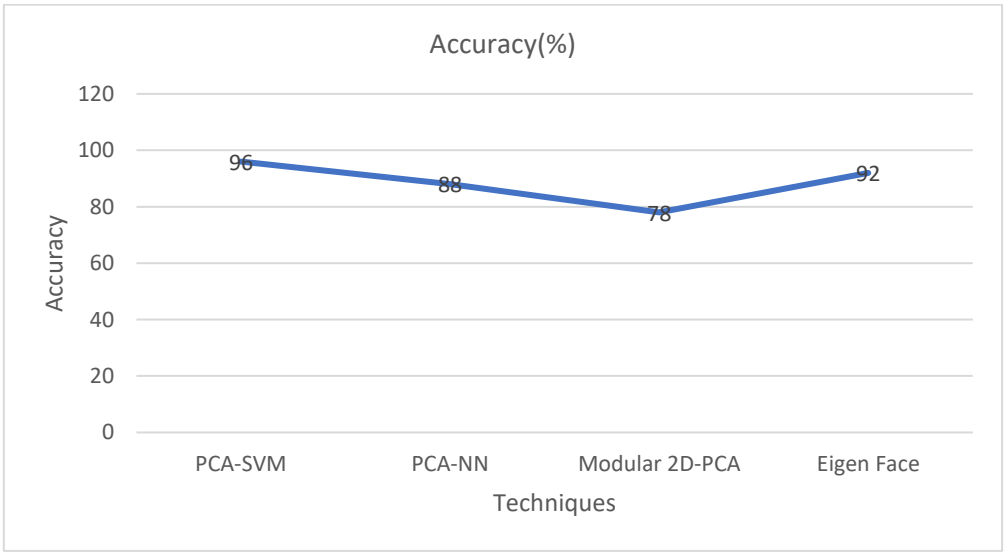

Figure 8: Graph representing Accuracy Rate

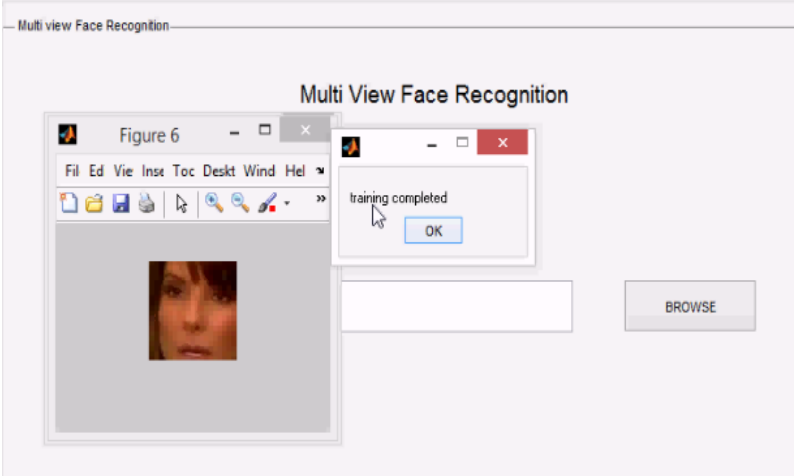

Figure 9: GUI Implementation of training dataset

Table 4: Results for matrix

Distance points	Distance range	Result
1	6.0617e+04 + 7.9017e-13i	-1.84
2	1.5465e+06 + 1.3990e-11i	+0.14
3	9.2378e+05 + 9.4015e-12i	-0.16
4	1.6832e+07 + 2.0898e-11i	-0.40
5	1.6129e+07 + 1.0569e-12i	+0.55
6	1.3844e+08 + 1.8304e-11i	-0.44

The results in table 4 show that the maximum distance result is +0.55 with maximum distance range's points 12i. The minimum distance result is -1.84 with maximum distance range's points 11i and 13i. The result of proposed distance learning process is much better and precise than the existing process.

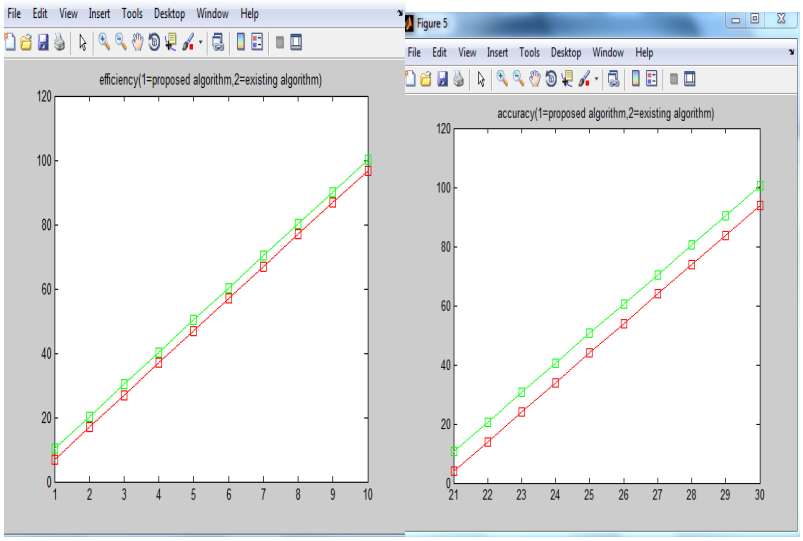

Figure 10: Comparison results between the proposed and existing algorithm

The results drawn from figure 10 showcases the efficiency of proposed vs existing methodologies. It can be seen that efficiency and accuracy increase along with novel points. The difference between both the approaches can be observed with the mentioned figure.

Amongst all the approaches PCA and SVM has been observed with best outcomes and efficiency. Following can be concluded from the study:

1. These results highlight how bad image quality affects face recognition algorithms' performance and point to the necessity of corrective measures.
2. The PCA model performed better for recognition. This confirmed the relative efficacy of using certain statistical imputation techniques in handling occlusions, as mentioned.
3. The suggested model demonstrated comparatively greater average recognition rates than some of the current improved methods.
4. It is suggested that PCA along with integration of SVM and CNN can result into better accuracy and future model can be developed with optimized approach and better integrations.

From the above results PCA is technique which helps in face recognition with better accuracy results and also keeps the important features and parameters intact. But the issue of Occlusion and pose is yet to be optimised for 100% results in accuracy and efficiency.

REFERENCES

[1] Rashmirekha Mohanty, Chandrakanti Malik , Gayatri Barik, Pramita Senapati4 , Zeenat Afroz, "A Comprehensive Review on Face Recognition System", International Journal of Creative Research Thoughts.
[2] H. L. Gururaj, B. C. Soundarya , S.Priya, J.Shreyas1, And Francesco Flammin, "A Comprehensive Review of Face Recognition Techniques, Trends, and Challenges, 2024
[3] S. V. Tathe, A. S. Narote, S. P. Narote, " Face Detection and Recognition in Videos," IEEE 2016.
[4] R. Sharma and M.S. Pattern, "A new pose invariant face recognition system using PCA and ANFIS," International Journal for Light and Electron Optics vol. 126, no. 23, pp. 3483-3487, December 2015
[5] Dhavalsinh V. Solanki *et al., "A survey on face recognition techniques and its applications, 2014.*
[6] Hamid Reza Yazdani , Ali Reza Shojaeifard, "Facial recognition system using eigenfaces and PCA", Computational and Applied Mathematics, March 2023.
[7] Zhao W, Chellappa R, Phillips PJ, Rosenfeld A. Face recognition: a literature survey. ACM computing surveys (CSUR). 2003; 35(4):399-458.
[8] Liau HF, Ang LM, Seng KP. A multiview face recognition system based on eigenface method. In 2008 International Symposium on Information Technology 2008 (pp. 1-5). IEEE
[9] Bidyut Jyoti Saha, Amitava Sen and Subnum Begum, " Face Pose estimation using PCA", ACCENTS Transactions on Image Processing and Computer Vision, Vol 2(4),2016.
[10] G. Guo, N. Zhang, A survey on deep learning-based face recognition, Comput. Vis. Image Underst. 189 (2019) 102805.
[11] John K. Esse, Joseph A. Mensah, Eric Ocran, Louis Asiedu, " On the search for efficient face recognition algorithm subject to multiple environmental constraints", Heliyon 10 (2024) e28568, 2024.
[12] D. Wang, S.C.H. Hoi, Y. He, and Jianke Zhu, "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation," IEEETransactions on Knowledge And Data Engineering, vol. 26, No. 1, January 2014
[13] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, Dealing with occlusions in face recognition by region-based fusion, in: Proceedings of the 2016 IEEE International Carnahan Conference on Security Technology (ICCST), 2016, pp. 1–6.
[14] Hamid Reza Yazdani, Ali Reza Shojaeifard , "Facial Recognition System using Eigenfaces and PCA", Computational and Applied Mathematics · March 2023.
[15] H. Moon, P.J. Philips, Computational and performance aspects of PCA - based face recognition algorithms, Perception, 30 2001, 303-321
[16] P. Kamencay, R. Hudec, M. Benco, M. Zachariasova, 2D-3D Face Recognition Method Based on a Modi_ed CCA-PCA Algorithm, International Journal of Advanced Robotics Systems, 2014, 215-230.

[17] V. Jalaja, G.S.G.N. Anjaneyulu, “Face RecognitionBy Using Eigen Face Method”, International Journal Of Scientific & Technology Research Volume 9, Issue 03, March 2020.
[18] http://cswww.essex.ac.uk/mv/allfaces/faces96.html
[19] Rahul Nandal , Dr. Anil Dudy, “Face Recognition Using SVM Based Machine Learning: A Review”, Webology, Volume 18, Number 6, 2021.
[20] Muge Carikcia , Figen Ozen, “A Face Recognition System Based on Eigenfaces Method”, Procedia Technology 1 (2012) 118 – 123
[21] M. Deshpande. Face recognition with eigenfaces – computer vision tutorial. Section: Computer Vision. [Online]. Available: https://gamedevacademy.org/ face-recognition-with-eigenfaces
[22] Imran M.A, Miah M. S. U, Rahman. H , Bhowmik A, Karmaker D ,] “Face Recognition Using Eigenfaces”, in International Journal of Computer Applications, 2015.
[23] Janarthany Nagendrajah, “Recognition of Expression Variant Faces- A Principle Component Analysis based on approach for access control, “ IEEE in International Conference On Information Theory and Information Security, 2010.
[24] Chung- Hua Chu, Yu-Kai Feng, “Study of Eye Blinking to Improve Face Recognition for Screen Unlock on Mobile Services”, in „ Journal of Electrical Engineering Technology“, 2017, pp. 1921- 1928.
[25] Yan. H, Wang. P, Chen. W.D, Liu. J, “Face Recognition Based on Global Wavelet Transform and Modular 2DPCA”, in International Conference on Power Electronics and Energy Engineering, 2015.
[26] Anjana Mall, Mrs. Shusmita Ghosh, “A Neural Network Based Face Detection Approach”, International Journal of Computer Technology and Applications, July 2010.
[27] Karagulle. F, “Face Finding Using Support Vector Machines”, M.S Thesis, Trakya university, Turkey, 2008.
[28] Alireza Tofigh, Nima Khairdoost, S. Amirhassan Monadjemi, Kamal Jamshidi,”A Robust Face Recognition System in Image and Video”, I.J. Image, Graphics and Signal Processing, 2014, 8, 1-11.
[29] Shrutika Yawale, Mayuri Patil , “Face Recognition using Eigenfaces for Android Application”, –International Journal for Innovative Research in Science & Technology| Volume 1 | Issue 6 | November 2014 ISSN (online): 2349-6010.
[30] Animasaun, N.A. Shah, A. Wakif, B. Mahanthesh, R. Sivaraj, O.K. Koríko, Ratio of Momentum Diffusivity to Thermal Diffusivity: Introduction I .L., Meta-Analysis, and Scrutinization, CRC Press, 2022.
[31] Joseph A. Mensah a, Justice K. Appati c, Elijah K.A Boateng a, Eric Ocran, “FaceNet recognition algorithm subject to multiple constraints, Scientific African 23 (2024) e02007.
[32] Rahul Nandal, Dr. Anil Dudy, “Face Recognition Using SVM Based Machine Learning: A Review”, Webology Volume 18, Number 6, 2021.
[33] Faridi, M. S., Zia, M. A., Javed, Z., Mumtaz, I., & Ali, S. (2021). A comparative analysis using different machine learning: An efficient approach for measuring accuracy of face recognition. International Journal of Machine Learning and Computing, 11(2)
[34] Tabassum, F., Islam, M. I., Khan, R. T., & Amin, M. R. (2020). Human face recognition with combination of DWT and machine learning. Journal of King Saud University-Computer and Information Sciences.