Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Enhancing Face Recognition through Dimensionality Reduction Techniques and Diverse Classifiers

Aziz Makandar[1], Shilpa Kaman[2], Syeda Bibi Javeriya[3]
[1]Professor, Dept. of Computer Science, KSAWU, Vijayapura, Karnataka, India.
[2,3]Research Scholar, Dept. of Computer Science, KSAWU, Vijayapura, Karnataka, India.
shilpaksawu@gmail.com, syedajaveriya84@gmail.com

DOI: 10.5281/zenodo.11109936

ABSTRACT

Face recognition is essential component of various applications including computer vision, security systems and biometrics. By examining the efficacy of several dimensionality reduction techniques, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Singular Value Decomposition (SVD), and Non-Negative Matrix Factorization (NMF), this paper offers a novel approach to face recognition. These techniques are combined with diverse classifiers, including Support Vector Machines (SVM), Random Forest, LightGBM, and k-Nearest Neighbors (KNN), are employed to evaluate their impact on face recognition accuracy. Experiments were conducted on Olivetti faces data set. We have demonstrated the comparative analysis of different dimensionality reduction techniques classifiers in terms of accuracy, precision, recall, f1score. Results shows the potential of integrating PCA with diverse classification models to enhance face recognition accuracy and highlights its applicability in real-world scenarios.

Keywords: Face recognition, PCA, Eigen faces, SVD, LDA, NMF, KNN, Random Forest, SVM, LightGBM

Cite as: Aziz Makandar, Shilpa Kaman, & Syeda Bibi Javeriya. (2024). Enhancing Face Recognition through Dimensionality Reduction Techniques and Diverse Classifiers. *LC International Journal of STEM, 5*(1), 36–44. **https://doi.org/10.5281/zenodo.11109936**

INTRODUCTION

In the fields of pattern recognition and image processing, face recognition is a difficult topic that presents challenges for both software and hardware, as the latter is developing algorithmic solutions while the former is producing physical implementations. [1][3]. The ability to accurately identify individuals and detect fraudulent attempts in a timely manner holds immense practical significance. There are several and diverse tracks in the planning of automatic face recognition algorithms. The task of determining the degree of similarity between a specific pair of facial images falls under this responsibility. As is evident in the subsequent fundamental stages, these tasks are carried out by advanced face recognition algorithms; they involve feature extraction and classification along with preprocessing of the facial image [2][3].

Principal Component Analysis (PCA) has emerged as a cornerstone technique in addressing these challenges, offering powerful tools for feature extraction, dimensionality reduction, and statistical modeling of facial data. PCA, originally introduced by Pearson [4] and further developed by Hotelling [5], has found widespread application in various domains due to its simplicity and effectiveness in capturing relevant information while reducing data dimensionality. Extraction of useful information from facial photographs is the aim of face analysis [6][7]. Technologically, Turk and Pentland [8] developed a near real-time face recognition system in 1991 by using PCA to calculate the Eigen face of

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

the images in the trained dataset. Although they were able to attain a respectable identification rate, their system is still susceptible to variations in orientation and lighting. The components of the Discrete Cosine Transform (DCT) and the outcomes of PCA were merged by Alex al et [9]. Poon al et al. [10] clipped the faces from the photos, used the Gradient Faces Algorithm to preprocess the facial pictures, and then used PCA to analyse the results. The illumination invariant features effect is eliminated. In their biometric identification system, Haghighat et al. [11] included a Gabor filter with PCA. Almotiri, et al [12] proposed face Recognition using PCA and Clustered Self-Organizing Map. A new combination of two face recognition algorithms based on projection It is suggested by [13] to use singular value decomposition (SVD) and relevance weighted linear discriminant analysis (RW-LDA) using the left and right singular vectors. The application of a well-known moment, Legendre, to a vector of features generated by the singular value decomposition transform (SVD) proposed in [14] led to the fusion of features. A method for extracting basic features from the face structure using Non-negative Matrix Factorization (NMF) is presented in the study [15]. Face recognition using different machine learning classifiers work is carried out by many researchers [16][17]. The proposed work is an attempt to study the effectiveness of different dimensionality reduction techniques combined with classifiers to recognize facial images.

METHODES

Dimensionality Reduction Techniques

In the discipline of machine learning, dimension reduction is an essential technique used to minimise the number of features being considered. By doing so, it seeks to simplify the complexity of datasets while retaining meaningful information. This procedure aids in resolving problems such as the dimensionality burden, which can cause overfitting in models and increased computational complexity. Common methods for dimensionality reduction include:

Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique used to condense large datasets while preserving essential information. Example In face recognition, it compresses pixel values, creating Eigen faces to represent unique features efficiently. Steps include establishing a training dataset, converting images to vectors, computing the mean image, subtracting it, performing SVD to find eigenvalues and eigenvectors, reconstructing images with the top eigenvalues, storing Eigen faces, establishing a testing dataset, projecting and subtracting test images, calculating eigenvalues/eigenvectors, then recognizing the closest match by calculating Euclidean distances [18]. In our experiment the number components obtained after applying dimensionality reduction with PCA are 150 as shown in the following figure 1 that contributes in increasing the performance of face recognition. Similarly, the top eigen faces obtained are represented in figure 2.

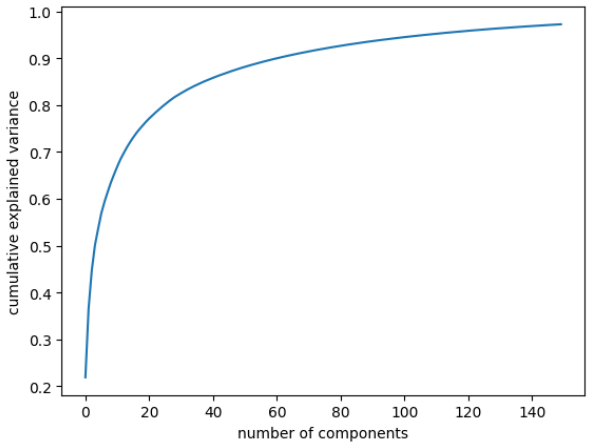

Figure -1: PCA Components Extracted From The Given Image.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

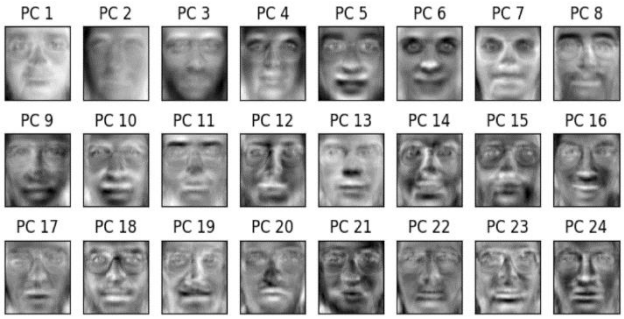

Figure -2: Eigen Faces After Applying PCA.

Linear Discriminant Analysis (LDA)

LDA is a classification algorithm as well as a method for reducing dimensionality. It looks for the linear feature combinations in the data that best distinguish between various classes or categories. Unlike PCA, which focuses on maximizing variance, LDA aims to maximize the separation between classes. In applications involving pattern recognition and classification, LDA is frequently utilised for feature extraction [19]. LDA is used to find the most discriminative features for face recognition. The number components we acquired after applying LDA to our data are 39, which is quite small in comparison to all the other dimensionality reduction approaches we are utilising. This intern shortens the model's execution time, which is further discussed in the results section that follows.

Singular Value Decomposition (SVD)

SVD is a method for matrix factorization that splits a matrix into three matrices: U, Σ, and V^T. The diagonal matrix Σ contains the singular values, while the other two matrices, V and U, are orthogonal. SVD is frequently used in the context of dimensionality reduction to reduce the dimensionality of data by keeping only the singular values and their accompanying singular vectors that are the most significant. Applications for SVD can be found in data analysis, recommendation systems, and picture compression [20]. 150 components were obtained following the application of SVD.

Non-negative Matrix Factorization (NMF)

A non-negative matrix can be broken down into two lower-dimensional matrices, one of which has only non-negative entries, using the dimensionality reduction technique known as NMF. It is particular useful for datasets where all values are non-negative, such as text documents, images, and audio signals. NMF has applications in topic modeling, image processing, and signal processing [21]. Number of components obtained are 150 after applying NMF and non-negative eigenfaces are shown in following figure 3.

Figure -3: Non-Negative Eigen Faces After Applying NMF.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Classifiers

Traditional machine learning methods including Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes, and KNN are used to perform face recognition. Each of these classifiers has its strengths and can contribute to improve the accuracy rate of face recognition.

Support Vector Machine (SVM)

Because Support Vector Machines (SVMs) are good at processing high-dimensional data and finding the best decision limits, they are frequently utilised in face recognition applications. A supervised machine learning approach that may be applied to regression and classification problems is called support vector machine. Finding the hyperplane that best divides the data points of several classes is the basic notion of SVM.

Random Forest (RF)

Another well-liked classifier for facial recognition is Random Forest. As part of its ensemble learning process, Random Forest builds a large number of decision trees during training and outputs a class that represents the mean prediction (regression) or mode of the classes (classification) of each individual tree.

LightGBM

LightGBM is a gradient boosting framework designed to handle huge datasets quickly and effectively. To lower computing costs during training, it uses strategies like Exclusive Feature Bundling (EFB) and Gradient-based One-Side Sampling (GOSS). LightGBM provides great performance and scalability using histogram-based decision trees and leaf-wise tree growth, which makes it a popular choice for a variety of machine learning tasks, especially when working with large-scale datasets.

KNN

A straightforward and user-friendly machine learning approach for classification and regression problems is called k-Nearest Neighbours, or kNN. In the context of face recognition, kNN can be employed to classify a given face into one of several predefined classes (individuals).

PROPOSED WORK

The working is being divided into two steps one is dimensionality reduction phase and classification phase. Following figure 1 shows the data flow diagram of the proposed methodology. The description of the dataset used in the work is provided below, along with a detailed step-by-step explanation of the procedure.

Step 1: Read the dataset, perform required preprocessing and split it into training and testing sets.

Step 2: Perform dimensionality reduction to extract features and reduce the number of components using following techniques.

- PCA is applied and the number of components is set to 150.
- Truncated SVD is applied and the number of components is set to 150.
- LDA is applied and the number of components is set to be the minimum of the unique target labels minus one and the number of features.
- NMF is applied and the 'random' initialization method is used, number of components is set to 150.

Step 3: Training data is projected to PCA, LDA, SVD and NMF components obtained after dimensionality reduction.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Step 4: Classifiers are initialized and trained on the transformed training data.

- SVM classifier with a radial basis function (RBF) kernel is initialized, Parameters such as C and gamma are set.
- Random Forest and lightGBM Classifiers are initialized with 100 trees.
- KNN classifier is initialized with the specified number of neighbors set to 5.

Step 5: Analyze how well the various models perform in terms of recall, accuracy, precision, and f1 score, also evaluate the execution times.

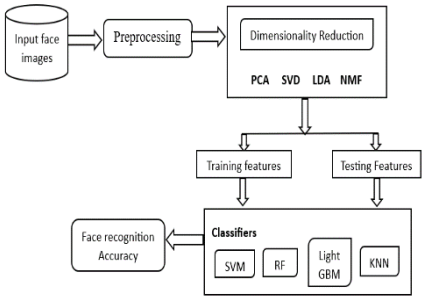

Figure -4: Workflow Diagram Of Face Recognition System.

Dataset

We are using Olivetti faces dataset [22][23] which is publicly available database for research purpose.

There are a total of 40 individuals (total 10 images per person) included in the dataset. The collection contains 400 face images. Face images were taken in various lighting conditions, with distinct expressions and levels of detail. Every facial image has a grayscale backdrop and a black level. Every image has a 64x64 size with pixel values scaled to an interval of [0, 1]. Following figure 2 shows sample input images.

Figure-5: Sample Input Images.

RESULTS AND DISCUSSION

Following table 1 gives the detailed performance analysis done using PCA, SVD, LDA and NMF dimensionality reduction techniques with various classifiers such as SVM, RF, LightGBM and KNN for face recognition. SVM is performing well with all the dimensionality reduction techniques when compared with other classifiers and at the same time LDA with random forest is very efficient in terms of accuracy. Results of LDA are good w.r.to all the classifiers compared with other three techniques.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

These results are represented graphically in figure 6 showing the performance of all the techniques used for face recognition.

Table-1: Face Recognition Results After Dimensionality Reduction With Various Classifiers.

	Accuracy	Precision	Recall	F1 score
PCA				
SVM	0.95	0.9675	0.95	0.9508
RF	0.92	0.94	0.92	0.917
LightGBM	0.75	0.869	0.75	0.749
KNN	0.76	0.846	0.76	0.779
SVD				
SVM	0.93	0.9455	0.93	0.9287
RF	0.92	0.9566	0.92	0.9271
LightGBM	0.75	0.7765	0.75	0.7344
KNN	0.79	0.7951	0.79	0.7729
LDA				
SVM	**0.97**	**0.9845**	**0.97**	**0.9728**
RF	**0.99**	**0.995**	**0.99**	**0.9909**
LightGBM	0.85	0.8753	0.85	0.8449
KNN	0.76	0.8517	0.76	0.7722
NMF				
SVM	0.91	0.92	0.91	0.9059
RF	0.83	0.8925	0.83	0.8293
LightGBM	0.75	0.816	0.75	0.7542
KNN	0.6	0.7153	0.6	0.5897

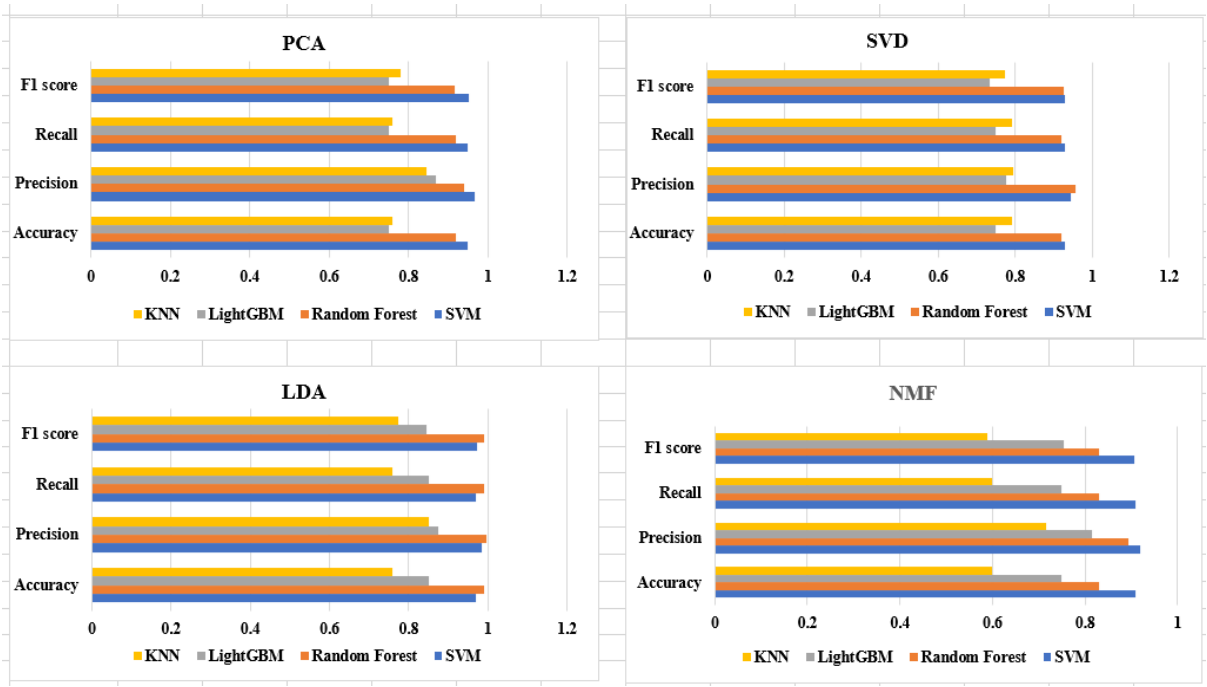

Figure-6: Accuracy Analysis of Different Classifiers on PCA, LDA, SVD and NMF For Face Recognition.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Accuracy comparison of random forest and SVM on different dimensionality reduction techniques such as PCA, SVD, LDA and NMF. We are comparing these two classifiers as their results are good compared with other classifiers. From the graph we can also say that the accuracy of LDA is good compared with other dimensionality reduction techniques used in the study. We are also comparing the time taken by different techniques as the number of components used by dimensionality reduction techniques are different which is shown in following table 2. It is being noticed that LDA is faster compared to remaining methods w.r.to all the classifiers.

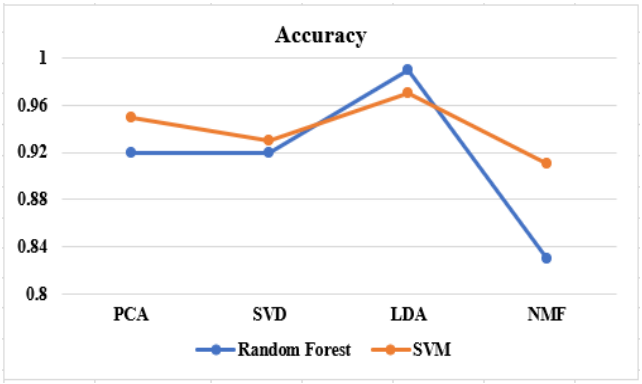

Figure -7: Accuracy Comparisons of RF and SVM.

CONCLUSION AND RECOMMENDATIONS

The paper presents the application of dimensionality reduction in conjunction with various classifiers for enhancing the performance of face recognition system. Use of dimensionality reduction techniques like PCA, SVD, LDA and NMF helps in extracting essential facial features, which is crucial for face recognition task. Furthermore, the integration of various classifiers such as SVM, RF, lightGBM and KNN contributes to the robustness of face recognition system. Overall experiments conducted shows that SVM and RF classifiers are performing well with all the dimensionality reduction techniques. If we compare efficiency of feature extraction techniques LDA is performing better with accuracies of 97% with SVM and 99% with RF classifiers. Execution time taken by LDA is also less compared to the remaining techniques used in the study. Hence, the combination of dimensionality reduction techniques and classifiers presents strong approach for face recognition, which can be adapted to different datasets to potentially improve the recognition accuracy.

Table-2: Comparison of Execution Times of Different Classifiers for PCA, LDA and SVD

Classifiers	Execution time in Seconds		
	PCA	SVD	LDA
SVM	4.885	5.245	3.115
Random Forest	5.211	7.427	3.851
LightGBM	91.056	158.648	88.366
KNN	3.072	2.812	2.821

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

REFERENCES

[1]Joshi, Asavari G. Deshpande, A.S. 2015. Review of Face Recognition Techniques. International Journal of Advanced in Computer Science and Software Engineering. 5(1). pp:71- 75.

[2] Klare, Brendan F. 2012. Heterogeneous Face Recognition. Ph.D. Thesis. Computer Science and Engineering. Michigan State University. East Lansing, Michigan, United States.

[3] Ibrahim, Fatima & Abdulmunim, Matheel. (2016). Design of an Efficient Face Recognition Algorithm based on Hybrid Method of Eigen Faces and Gabor Filter. Iraqi Journal of Science. 57. 2102-2110.

[4] Karl Pearson F.R.S. (1901) LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2:11, 559-572, DOI: 10.1080/14786440109462720.

[5] Hotelling, Harold. "Analysis of a complex of statistical variables into principal components." Journal of Educational Psychology 24 (1933): 498-520.

[6] Gross, R., Baker, S., Matthews, I., Kanade, T. (2005). Face Recognition Across Pose and Illumination. In: Handbook of Face Recognition. Springer, New York, NY. https://doi.org/10.1007/0-387-27257-7_10.

[7] Shakhnarovich, G., Moghaddam, B. (2005). Face Recognition in Subspaces. In: Handbook of Face Recognition. Springer, New York, NY. https://doi.org/10.1007/0-387-27257-7_8.

[8] Turk,M. Pentland,A. 1991. Eigenfaces for recognition. Journal of Cognitive Neuroscience. 3(1), pp:71-86.

[9] Alex, N.Sare. Reddy, Decpthi M. Reddy, Devika. 2016. A Study on Biometric Face Recognition for Login. International Journal of Innovative Research in Computer and Communication Engineering. 4(3). pp: 3096-3102.

[10] Poon, Buruce. Amin, Ashraful. Yan, Hong. 2016. Improved Methods on PCA Based Human Face Recognition for Distorted Images. Proceedings of the International MultiConference of Engineers and Computer Scientists. Hong Kong, March 16 – 18.

[11] Haghight, Mohammad. Zonouz, Saman. Abdel-Motaleb, Mohamed. 2015. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. Expert Systems with Applications. 42(21). pp: 7905-7916.

[12] Almotiri, Jasem. "Face Recognition using Principal Component Analysis and Clustered Self-Organizing Map." International Journal of Advanced Computer Science and Applications 13.3 (2022).

[13] Maafiri Ayyad, Chougdali Khalid, New fusion of SVD and Relevance Weighted LDA for face recognition, Procedia Computer Science, Volume 148, 2019, pp: 380-388, ISSN 1877-0509.

[14] Hashimi, Asaad Noori, and Buraq Noaman Kadhim. "Face recognition based on fusion of SVD and Legendre moment." In Journal of Physics: Conference Series, vol. 1530, no. 1, p. 012120. IOP Publishing, 2020.

[15] Nikan, F. and Hassanpour, H., 2020. Face recognition using non-negative matrix factorization with a single sample per person in a large database. Multimedia Tools and Applications, 79(37-38), pp.28265-28276.

[16] Parveen, P. and Thuraisingham, B., 2006, November. Face recognition using multiple classifiers. In 2006 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06) (pp. 179-186). IEEE.

[17] D. Sudiana, M. Rizkinia and F. Alamsyah, "Performance Evaluation of Machine Learning Classifiers for Face Recognition," 2021 17th International Conference on Quality in Research (QIR): International Symposium on Electrical and Computer Engineering, Depok, Indonesia, 2021, pp. 71-75, doi: 10.1109/QIR54354.2021.9716171.

[18] Adeshara, Kavan & Elangovan, Vinayak. (2020). Face Recognition using PCA Integrated with Delaunay Triangulation. 321-331. 10.5121/csit.2020.101424.

[19] F. Mahmud, M. T. Khatun, S. T. Zuhori, S. Afroge, M. Aktar and B. Pal, "Face recognition using Principal Component Analysis and Linear Discriminant Analysis," 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Savar,

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Bangladesh, 2015, pp. 1-4, doi: 10.1109/ICEEICT.2015.7307518
[20] Zeng, G. (2007). Facial Recognition with Singular Value Decomposition. In: Elleithy, K. (eds) Advances and Innovations in Systems, Computing Sciences and Software Engineering. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6264-3_26
[21] W.-S. Chen, J. Liu, B. Pan, and B. Chen, "Face recognition using nonnegative matrix factorization with fractional power inner product kernel," Neurocomputing, vol. 348, pp. 40-53, 2019, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2018.06.083.
[22] Liu, Nelson (2016). scikit-learn olivetti faces dataset olivettifaces.mat. figshare. Dataset. https://doi.org/10.6084/m9.figshare.3829989.v2

AUTHORS PROFILE

Dr. Aziz Makandar currently working as a Professor in the Department of Computer Science, Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India. His qualifications are Ph.D (Computer Science), M.Tech. and B.E. (Computer Science and Engineering). Having more than of 21 years of teaching and research experience he supervised many PhD and MPhil holders. His Areas of Research Interests includes Digital Image Processing, Pattern Recognition, Knowledge Acquisition, Artificial Intelligence, Machine Learning, Deep learning, Artificial Neural Network. He has published more than 130 papers comprising National, International conferences and Journal publications with 830 citations (14, 17 h index and i10 index respectively).

Mrs. Shilpa Kaman is a research scholar at the Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India. Her qualifications are M.Phil. (Computer Science), M.Tech. and B.E. (Computer Science and Engineering). She has 9 years of teaching and research experience and her areas of interest include Digital Image Processing, Artificial Intelligence, Machine Learning, Deep Learning, Image forensics.

Miss. Syeda Bibi Javeriya is a research scholar at the Karnataka State Akkamahadevi Women's University, Vijayapura, Karnataka, India. Her qualifications are as mentioned M.Phil. (Computer Science), M.Sc. and B.Sc. (Computer Science and Engineering). She has 4 years of research experience and her areas of interest include Digital Image Processing, Artificial Intelligence, Machine Learning, Deep Learning, Image forensics.