iConclusions https://iconclusions.com/ Conclusions in Engineering

A Comparative Analysis for the Performance of LFW and ORL Databases in Facial Recognition

Mohamed Mosadag,[1, a)] **Mohammed Ahmed Mohammed**,[1] **Kamal Bashir**,[1] and **Amin Mubark Alamin**[2]

[1)] *College of Computer Science and Information Technology, Karary University, Omdurman 12304, Sudan*

[2)] *College of Computer Science and Information Technology, Alneelain University, Khartoum, Sudan*

ABSTRACT: It has become clear in the last many years that face recognition has attracted the attention of researchers. Until now, this has been a difficult area of research due to several problems. The face continues to be the most difficult subject of study for experts within the realm of computer vision and image processing, because it is an element with diverse sensory properties. There are several face images databases used to train and test facial recognition systems used in the recent literature with different results. But until now, the most effective database has not been identified. The free facial image databases LFW (Labeled Faces in the Wild) and ORL (Olivetti Research Laboratory) are the most widely used. The main idea of this survey work is to compare the selected face databases based on using different face recognition methods (different feature extraction and classification techniques).The performance, accuracy, and computational requirements of these methods are analyzed through a series of case studies and empirical evaluations with a confusion matrix. The results of this comparison will be used in the author's future work in the face detection and recognition software environment for evaluating performance and testing purposes. The results obtained support that the suitable choice between LFW and ORL databases depends on the specific goals of the facial recognition system. For initial development and controlled testing, ORL is highly effective. For evaluating performance in real-world scenarios, LFW is more representative but challenging.

Received: 14 January 2025
Accepted: 13 February 2025
DOI: https://doi.org/10.71107/pydt9t88

I. INTRODUCTION

In recent years, the contemporary world has evolved into a globalized entity. The frightening and frustrating reality is that some of the benefits of technology are being used negatively to put people and their property at risk. In line with this continuous technological progress and development, the need to take precautions to protect people and their rights is gradually increasing. Due to the drawbacks of conventional methods like personal identification numbers, identity verification badges, and passwords, interest in biometric systems is growing. Among the most used biometric methods, facial recognition has become a central issue, and for a variety of explanations, the face has drawn several researchers. Originally, a human facial recognition system is not parasitical. Secondly, the development in digital cameras, storage media, and technology has made it possible to handle enormous face databases. A facial recognition system is a technology potentially capable of matching a human face from a digital image or video image to a database of faces[1].

Such a system is typically used to authenticate users through identity verification services and works by distinguishing and measuring facial features from a given image. Since computerized facial recognition entails the quantification of physiological attributes of individuals, facial recognition systems are classified as biometric modalities. Despite the fact that the precision of facial recognition systems as a biometric technology is inferior to that of iris recognition, fingerprint capture, palm identification, or voice recognition, it has garnered

a) Electronic mail: m.mosadag@gmail.com

extensive acceptance owing to its non-contact methodology. Facial recognition systems have been implemented in advanced human-machine interaction, video surveillance, law enforcement, passenger screening, employment and housing decisions, and automatic image indexing[2]. The recent progression of facial recognition technologies can be predominantly attributed to substantial developments in ML, coupled with sophisticated methods for analyzing data. Therefore, facial recognition has become a trustworthy technology for confirming one's identity. It is crucial to carry out a comparison analysis in order to assess the effectiveness of these various facial technologies, something that can only be done by conducting training and testing operations on a data set taken from the real world, by which we mean here in our study databases for facial images. Therefore, the primary goal of this paper was to conduct a comparison to distinguish the best between two of the most widely used free databases for training and testing recognition systems in order to provide ease and validity of choice for the researchers and those working on recognition systems. Because it was very important to correctly choose the best database on which the recognition system would be trained and tested.

II. BACKGROUND AND MOTIVATION

Face databases are used to test face detection and recognition algorithms. Various algorithms are capable of being applied to identical datasets, after which the outcomes are subjected to rigorous evaluation. There are also specialized face databases, e.g., to test invariant face detection and recognition methods, invariant methods for the individual aging problem, facial expression detection databases, and others.

The research conducted by Vaishali and Pramod Patel in 2015[3] elucidated the application of random forests in their methodology. They employed a synergistic approach involving Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), and Discrete Wavelet Transform (DWT) for the ORL database's preparation and vector feature extraction. A subsequent investigation utilizing Support Vector Machine (SVM) as a method of classification, in conjunction with PCA as an extractor of characteristics, achieved a commendable rate of recognition[4]. Other studies[5] yielded favorable outcomes by deploying Naive Bayes Classifiers integrated with PCA on the Yale dataset, contrasting with the application of Linear Discriminant Analysis (LDA) as a feature extractor. In 2016, various methodologies[6] were proposed to synergistically integrate PCA and LDA for feature extraction alongside Artificial Neural Networks (ANN) for classification purposes. Prior to these processes, image preprocessing techniques were employed, including histogram equalization, normalization of image dimensions, and the transformation of red, green, and blue (RGB) color photos into grayscale formats. The assessment of these systems resulted in favorable recognition rates based on the ORL database. In[7], the authors executed facial recognition tasks utilizing SVM, PCA, and Local Binary Patterning (LBP) to enhance accomplishment metrics. For collecting data, the Yale University Library (Yale) and ORL databases were selected. The area of the face in each picture was isolated employing the Viola-Jones algorithm, followed by resizing and cropping into 70x70 pixel dimensions. The authors applied contrast enhancement techniques to improve image quality. The feature extraction process was conducted separately utilizing both LBP and PCA. For categorization purposes, SVM was employed at construct paradigms, which were subsequently evaluated on the Yale and ORL databases. In[8], Bala et al. conducted a study on facial recognition employing K-nearest neighbors (KNN) classifiers. They devised the extraction phase utilizing Linear Discriminant Analysis (LDA). Huda et al. in[9] introduced a methodology employing the Random Forest (RF) classifier. Initially, the system employed the Viola-Jones algorithm for face detection within images. Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG) descriptors were concurrently applied to extract feature vectors. Classification was executed utilizing the Random Forest (RF) classifier. This system used Mediu-S-DB (Mediu staff database) to achieve commendable recognition rates.

In the year 2017, Bekhouche and Salah Eddine introduced an innovative learning framework aimed at the estimation of human demographic characteristics, wherein the attributes of ethnicity, gender, and age are deduced from facial imagery. Empirical investigations were conducted utilizing five publicly available databases (MORPH II, PAL, IoG, LFW, and FERET), alongside an additional two challenge datasets[10]. Lahaw et al. in[11] unveiled methodologies that incorporated PCA, LDA, Independent Component Analysis (ICA), and Discrete Wavelet Transform (DWT). The Two-Dimensional Principal Component Analysis (2D-DWT) serves as a multi-level decomposition technique utilized for the preprocessing of pictures. 4 separate bands are created from the primary image: (LL) low-low, (LH) low-high, (HL) high-low, and (HH) high-high. The separated band (LL) serves as the foundational inputting image for the extraction of the features procedure employing ICA, PCA, and LDA algorithms. The resultant features were subsequently categorized

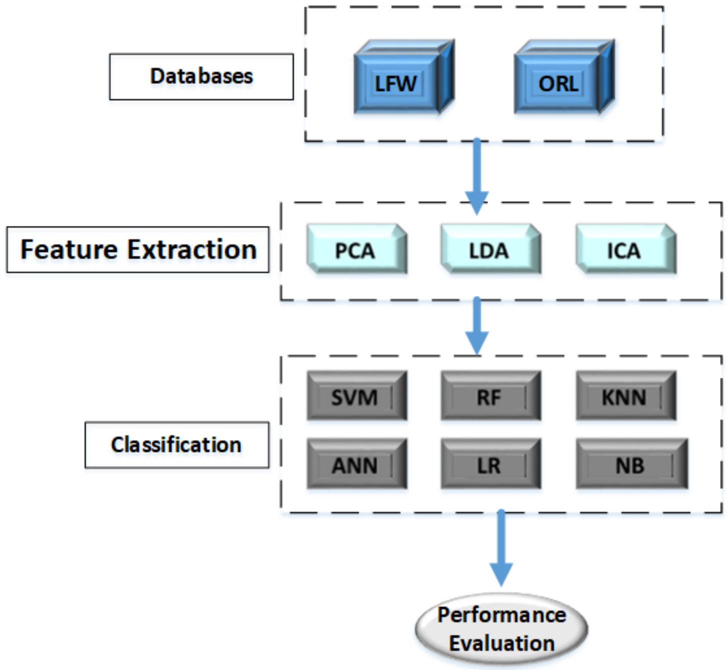

FIG. 1: Proposed models of facial recognition.

using SVM. The models (DWT \LDA\$SVM$), (DWT \PCA\$SVM$), and ($DWT\backslash ICA\backslash SVM$) achieved commendable recognition rates on the ORL database. In 2018, Pedro, Firas, and Shakir implemented DWT in conjunction with PCA techniques for the purposes of preprocessing and feature extraction, employing the KNN classifier for the execution of classification tasks[12]. They reported a favorable recognition rate on the ORL database.

In the year 2019, Sri et al.[13] and Kadek et al.[14] implemented two-dimensional Principal Component Analysis (2D-PCA) for facial representation. Furthermore, grayscale conversion, Haar cascade segmentation, and ROI (region of interest) delineation were employed by Ni Kadek et al. for image preparation. Any classifications were performed utilizing the KNN methodology on the ORL database.

In 2021, Ajeet Singh, Atul Pratap Singh, and Neha Verma engaged in a thorough examination of facial detection methodologies employing Artificial Intelligence (AI), with a particular emphasis on the amalgamation of PCA and KNN algorithms. In order to efficiently extract prominent features and minimize data loss, PCA was used to reduce the dimensionality of facial picture datasets. The KNN classifier was employed for categorization by identifying the nearest corresponding face within a dataset. By implementing these methodologies on the LFW dataset, an overall accuracy rate of 88% was attained[15].

Several face databases were selected for further testing of the software environment, developed by the authors, which will be used for further evaluation and improvement of face detection and recognition algorithms. The software environment is in an early stage of development:

A. Labeled Faces in the Wild (LFW)

It is an image database containing pictures of faces designed to evaluate face recognition algorithms in more realistic and uncontrolled environments. Collected from the web specifically to study the problem of unrestricted facial recognition[16]. It contains a large number of images (13,000 images of 5,749 subjects collected from the web, with variations in lighting, exposure, and background). The LFW database was developed and maintained by researchers at the University of Massachusetts, Amherst. It was released for research purposes to advance facial verification. The original database contained four different sets of LFW images as well as three different types of "aligned" images suitable for testing the robustness and performance of facial recognition systems in real-world scenarios.

B. Olivetti Research Laboratory (ORL)

Contains 400 images of 40 separate subjects (10 images per person) with different lighting, facial expressions, and facial details. Presented by Samaria and Harter in parameterizing a stochastic model for human face identification. It was collected from 1992 to 1994 in the laboratory[17]. It is mainly used for controlled experiments in face recognition. Ideal for testing algorithms under controlled conditions with changes in pose and expression.

III. THE PROPOSED FACIAL RECOGNITION MODEL

The processes of feature extraction and classification represent two fundamental components within facial recognition systems. This study will undertake a comparative analysis of various selected efficient techniques that are frequently employed in these phases. Figure 1 illustrates the methodologies proposed to implement our model for the facial recognition system, encompassing data acquisition (databases), feature extraction, classification, and subsequent model evaluation[18].

The methodologies employed in the feature extraction phase encompass PCA (Principal Component Analysis), which helps in reducing dimensionality and noise, simplifying the data for easier management and analysis; LDA (Linear Discriminant Analysis), which focuses on enhancing class separatability for improved recognition accuracy; and ICA (Independent Component Analysis), which works on capturing independent features, making the system more robust to variations. These are the three most frequently used techniques, and each one of them has its strengths. A plethora of machine learning techniques has been leveraged as classification methods for the purpose of face recognition. We utilize several prevalent supervised learning methodologies, specifically KNN (K-nearest neighbors), NB (Naive Bayes), SVM (Support Vector Machine), Random Forest (RF), MLP (Multi-Layer Perceptron) as a form of ANN (Artificial Neural Network), and LR (Logistic Regression).

The confusion matrix concerning the binary classification problem will be utilized to provide more clarity regarding accuracy metrics employed for evaluation purposes, presented in Table I. The true class label is shown in the first column, while the predicted class label is shown in the second and third columns. True Positive (TP) and True Negative (TN) indicate the number of correctly classified positive and negative samples, while False Negative (FN) and False Positive (FP) indicate the number of incorrectly classified positive and negative samples, respectively.

The metric of binary accuracy can be delineated in the following manner:

$$(\mathrm{TP}+\mathrm{TN})/(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})$$

IV. RESULTS AND DISCUSSION

Accuracy results and a performance comparison of facial image databases based on the proposed methods have been presented in this section. To execute the experiments in this study, a good computer and updated programming software were used. The overall specification was as follows: Laptop Specifications: HP notebook, Windows 10, processor (CPU) Intel Core i5 and 5200 U 5th Generation, system RAM 2 GB, storage 1 TB HDD, camera front. As a software language, Python 3.9.0 was used to apply our algorithms and conduct the whole experiment.

Overall, as shown in Table II, the performances of the LFW database, with most methods, are satisfactory. Where we see that (ICA+SVM) performs better worldwide. An average score of 83.54% was the outcome. With an average of 80.88%, ICA+MLP come right after it, followed by PCA+MLP with an average of 80.36%. The performance of PCA+KNN was the lowest in this work, at 42.12%.

Table III summarizes the performances of the LFW database. With the ORL database, facial recognition techniques work incredibly well overall. On a worldwide scale, LDA+LR performs better. An average score of 98.75% was the outcome. With an average of 97.50%, PCA+SVM and PCA+LR come right after it, followed by LDA+KNN with an average of 96.25%. The task with the lowest performance, ICA+NB, was 73.75%. Based on the findings compiled in Table II, Table III, and Figure 2, we clearly show all the methods of facial recognition score high accuracy results in the case of the ORL database than the LFW database.

V. CONCLUSION

These previous results of comparison prove that the ORL database has resulted in better results than LFW, but at the same time we cannot say ORL is more effective and efficient than LFW for evaluating and testing facial recognition systems. The clear differences in accuracy results between the LFW and ORL databases can be attributed to several main factors: data distribution, sample size, image variability, and different dimensions.

LFW has a large number of images (over 13,000) and an irregular distribution of samples per subject (some subjects may have an image, while others might have

TABLE I: Confusion Matrix for a Two-Class Problem.

Class	**Predicted as positive**	**Predicted as negative**
Actual positive class	True Positive (TP)	False Negative (FN)
Actual negative class	False Positive (FP)	True Negative (TN)

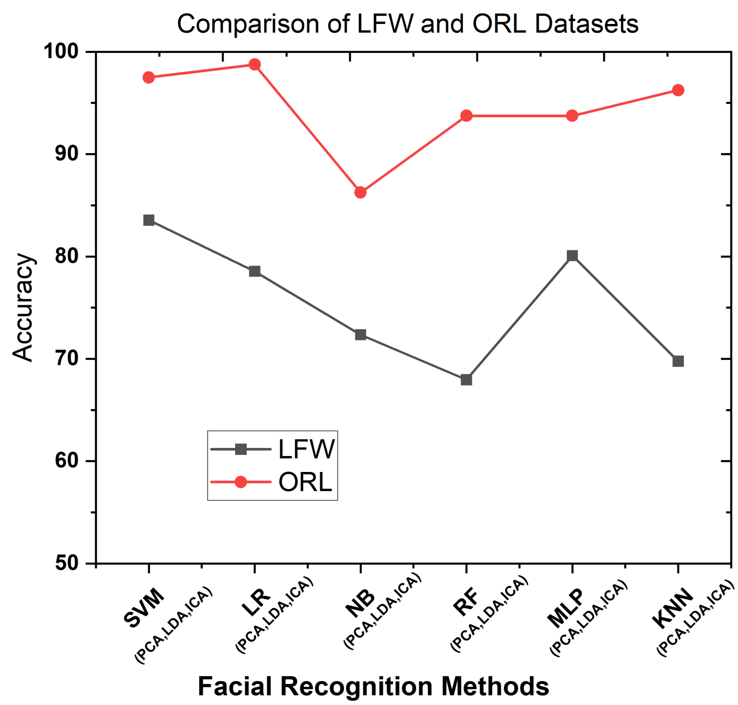

FIG. 2: Comparative performances of LFW database Vs. ORL database.

TABLE II: The results of different methods based on LFW database.

Methods	**ICA**	**LDA**	**PCA**
SVM	50.62	70.81	83.54
KNN	42.12	69.77	64.34
RF	57.88	67.96	57.11
LR	78.55	71.58	76.74
MLP	80.36	70.03	80.88
NB	72.35	70.54	70.03

TABLE III: The results of different methods based on ORL database.

Methods	**ICA**	**LDA**	**PCA**
SVM	97.50	93.75	93.75
KNN	87.50	96.25	75.00
RF	91.25	91.25	93.75
LR	97.50	98.75	93.75
MLP	85.00	93.75	91.25
NB	86.25	85.00	73.75

only one). This inconsistency can make it difficult for the model training to learn a large number of generalizable features; this leads to hindered performance. On the other hand, we find that ORL contains a smaller number of images (400 in total), but the uniform and balanced distribution of 10 images per subject makes it easier for the model to learn consistent features across different subjects; this helps to improve performance in terms of recognition accuracy. LFW has a high variability of images (variations in lighting, pose, expressions, and backgrounds) captured in uncontrolled environments; this variability makes it challenging for the model to extract consistent features. Also, the images are not standardized in terms of dimensions, which can introduce additional complexity in the preprocessing stage. On the other hand, we find that ORL images have minimal variability and are taken in a controlled environment; this consistency makes it easier for the model to extract relevant features. Also, all images are of the same size (92x112 pixels), simplifying the preprocessing and feature extraction process.

After deep discussion, it is concluded to say LFW is reflecting real-world scenarios (more representatives of actual use cases in facial recognition systems). So we can say it is suitable for benchmarking and real-world applications (unconstrained environments), but the irregular distribution and different dimensions can hinder performance. On the other hand, ORL offers a controlled and consistent dataset with regular distribution and standardized dimensions, leading to better performance in terms of recognition accuracy, but it may not fully represent real-world conditions (controlled settings). So we can use it in the initial stages of developing and testing facial recognition models due to its consistency and ease of use. Finally, this paper has investigated the suitable choice between LFW and ORL databases, which depends on the specific goals of our facial recognition system. For initial development and controlled testing, ORL is highly effective. For evaluating performance in real-world scenarios, LFW is more representative but challenging.

In future works we should explore the inclusion of more diverse and larger-scale databases to involve curated collections that encompass a wider range of ethnicities, ages, and environmental conditions to improve the generalizability of it in the processes of evaluating all face recognition systems.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] A. A. Raj, M. Shoheb, K. Arvind, and K. S. Chethan, "Face recognition based smart attendance system," in *2020 International Conference on Intelligent Engineering and Management (ICIEM)* (IEEE, 2020) pp. 354–357.

[2] S. K. Chen and Y. H. Chang, "Study of the effectiveness of various feature extractors for human face recognition for low resolution images," in *2014 International Conference on Artificial Intelligence and Software Engineering (AISE2014)* (DEStech Publications, Inc, 2014) pp. 1–7.

[3] V. W. Parate and P. Patel, "Pca, dct and dwt based face recognition system using random forest classifier," International Journal of Digital Application and Contemporary Research **3**, 1–7 (2015).

[4] P. S. Gharamaleki and H. Seyedarabi, "Face recognition using eigen faces, pca and support vector machines," European Journal of Applied Engineering and Scientific Research **4**, 55–63 (2015).

[5] T. A. H. and D. Sachin, "Dimensionality reduction and classification through pca and lda," International Journal of Computer Applications **122**, 100–104 (2015).

[6] G. Kaur and H. Kaur, "Efficient facial recognition using pca lda combination feature extraction with ann classification," International Journal of Advanced Research in Computer Science and Software Engineering **6**, 258–263 (2016).

[7] B. Pathya and S. Nainan, "Performance evaluation of face recognition using lbp, pca and svm," SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE) **3**, 85–88 (2016).

[8] M. Bala, P. Singh, and M. S. Meena, "Face recognition using linear discriminant analysis," International Journal of Electrical and Electronics Research **59**, 72–87 (2016).

[9] M. Huda and S. M. S. Hilles, "Face recognition and detection using random forest and combination of lbp and hog features," in *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)* (IEEE, 2018) pp. 1–7.

[10] S. E. Bekhouche *et al.*, "Pyramid multi-level features for facial demographic estimation," Expert Systems with Applications **80**, 279–310 (2017).

[11] Vanlalhruaia, Y. K. Singh, and N. Singh, "Binary face image recognition using logistic regression and neural network," in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)* (IEEE, 2017) pp. 3883–3888.

[12] S. F. Kak, F. M. Mustafa, and P. R. Valente, "Discrete wavelet transform with eigenface to enhance face recognition rate," Academic Journal of Nawroz University **7**, 9–17 (2018).

[13] S. Sutarti, A. Putra, and E. Sugiharti, "Comparison of pca and 2dpca accuracy with k-nearest neighbor classification in face image recognition," Scientific Journal of Informatics **6**, 64–72 (2019).

[14] N. K. A. Wirdiani, P. Hridayami, N. P. A. Widiari, K. D. Rismawan, P. B. Candradinata, and I. P. D. Jayantha, "Face identification based on k-nearest neighbor," Scientific Journal of Informatics **6**, 150–159 (2019).

[15] N. Singhal, V. Ganganwar, M. Yadav, A. Chauhan, M. Jakhar, and K. Sharma, "Comparative study of machine learning and deep learning algorithm for face recognition," Jordanian Journal of Computers and Information Technology **7** (2021), 10.5455/jjcit.71-1624859356.

[16] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition* (2007).

[17] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision* (1994) pp. 138–142.

[18] A. Dirin, N. Delbiaggio, and J. Kauttonen, "Comparisons of facial recognition algorithms through a case study application," International Journal of Interactive Mobile Technologies (iJIM) **14**, 121–133 (2020).