Face Recognition Comparative Analysis Using Different Machine Learning Approaches

Nisar Ahmed[1*], Farhan Ajmal Khan[2], Zain Ullah[3],
Hasnain Ahmed[1], Taimur Shahzad[4], Nableela Ali[5]

[1] Department of Computer Science, Sapienza University of Rome, Italy

[2] Department of Artificial Intelligence, Sapienza University of Rome, Italy

[3] Department of Data Science, Sapienza University of Rome, Italy

[4] Department of Computer Science, COMSATS University Islamabad, Attock Campus, Pakistan

[5] Department of Information Technology, Quaid-e-Awam University, Nawabshah, Pakistan

[*] Corresponding author's email: ahmed.1772799@studenti.uniroma1.it

ABSTRACT

The problem of a facial biometrics system was discussed in this research, in which different classifiers were used within the framework of face recognition. Different similarity measures exist to solve the performance of facial recognition problems. Here, four machine learning approaches were considered, namely, K-nearest neighbor (KNN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Principal Component Analysis (PCA). The usefulness of multiple classification systems was also seen and evaluated in terms of their ability to correctly classify a face. A combination of multiple algorithms such as PCA+1NN, LDA+1NN, PCA+ LDA+1NN, SVM, and SVM+PCA was used. All of them performed with exceptional values of above 90% but PCA+LDA+1N scored the highest average accuracy, i.e. 98%.

Keywords: linear discriminant analysis, k-nearest neighbor, support vector machine, principal component analysis.

INTRODUCTION

In recent times, the approaches that are based on the biometrics system have been materialized for the purpose of recognizing individuals. The approaches monitor the individual's rational and behavioral features in order to specify that particular identity for the authentication process and to provide them with the access to the database by entering passwords, security codes, smart access cards, visa, etc.; however, these kinds of technologies cannot be secured by passcodes and the pin codes can be stolen, disremembered, or guesstimated easily [1]. On the other hand, smart access cards can be lost, or can become hard to read for the systems. Nevertheless, biometric technologies of individuals cannot be mislaid, copied, or counterfeited.

Face recognition is the most authentic and less obtrusive technique that can be very swift if we compare it to other technologies such as finger and signature recognitions. Despite the swift advancement in image processing techniques, there are still uncontrollable issues in face recognition. The decision can be made on fundamental concepts, for example, the distances or similarities towards the reference patterns. According to the problem of face detection and face recognition, the problems of face recognition are frequently confused.

The goal of this research was to choose the best possible approach to the identification of images from a collection of approaches selected. The images based on the open-source face image dataset are recognized by each algorithm. The algorithm processes all images and transforms them into the appropriate formats and then acknowledges the images, which are subsequently evaluated by comparing the results with the original data set.

Independent factor analysis based on face recognition

Many face recognition algorithms are currently working on face representation that operates on unsupervised statistical approaches to machine learning. These techniques find a set of foundation images and define the faces as a linear combination of those images [2].

Principal Component Analysis (PCA) is a well-known example of such methods. The foundation images identified by PCA are most reliant on the pair relationships in the image database between pixels. In particular, for example, face recognition in this particular work, where the key records can be stored in high-order relationships between pixels, it seems logical to expect better foundation images to be determined by the methods responsive to these excessive order statistics. Independent Component Analysis, a generalization of PCA [3], is one such method.

Maintaining the Integrity of the Specifications

Eigen space-based face recognition is definitely one of the most successful methodologies in digital images for the computational recognition of faces [4]. Many Eigenspace-based facial recognition approaches have been proposed in conjunction with the Eigenface Algorithm. This research aims to provide independent comparative analysis between a variety of primary methods focused on Eigenspace. We assume that these partial studies are important, because comparisons are typically carried out by using the study categories of the implementations proposed by each methodology, which does not completely observe the matched working conditions for the algorithms. Quite commonly, a completion is carried out between the research capabilities as the contrast between techniques is instead carried out.

Elastic Bunch Graph Matching

One of the well-known techniques proposed for face recognition is elastic bunch graph matching. We propose various extensions to Elastic Bunch Graph Matching and its latest Landmark Model Matching variation in the achievements of this paper. In order to find the landmarks (facial feature factors), Landmark Model Matching, primarily based entirely on Gabor wavelets for characteristic extraction, is employed [5].

We show that the combination of grey-level profiles with Gabor wavelet features for feature extraction can be improved. In addition, with the beneficial resource of hybridizing Gabor wavelet, we achieve enhanced recognition rates with the Eigenface characteristics defined through Principal Component Analysis, which can provide the knowledge keeping in the entire face look [5, 6].

FACE RECOGNITION APPROACHES

There are many crucial techniques to the face recognition problem. These include Template matching, Fisher-faces, Feature-based, Appearance-based, and Mixed-based, etc. some of them are explained below. several distinct algorithms related to face recognition were created.

Feature-based

In this group, the feature-based approach, the faces are fragmented into different pieces that can be evaluated separately. Normally in this section, it comprehends the identification procedure of local features of the face, for example, nose, mouth, and eyes [21], In many different ways, this type of features could be analyzed, by implementing the simplest computation techniques or more innovative template matching. In early times, these approaches used a collection of geometric feature points to compute and compare the distances and angles accordingly to an actual face. In 2002, a researcher used twenty-eight different points of the face from the manually selected two-dimensional images, to characterize the face [22]. The method was reported accurately by their results for face recognition that all belong to the same pose positionings [7].

Appearance-based

There are many machine learning approaches and algorithms, as it is indicated above. We are able to use any algorithm for the process of object recognition. Many of them, however, are helpful and efficient for the use of the Fisher face algorithm, such as Principal Component Analysis using Eigenfaces, Linear Discriminant Analysis, and Elastic Bunch Graph.

Algorithms

The application of different machine learning approaches to the problems in face recognition

was investigated. For supervised learning face recognition methods (classifiers), the following were considered: Linear discriminant analysis 1-nearest Neighbor, Support Vector Machine, and Principal Component Analysis. The overall performance of several more than one classifier structure was also tested and evaluated in terms of their capability to properly classifying. It is also possible to design the template-based methods using statistical approaches such as SVM, PCA, LDA, Kernel approaches or Trace Transforms. In turn the approaches which are designed based on geometric functionalities are feature-based techniques that analyze adjacent facial-features and their geometric relationships. It is likewise referred to as a feature-based approach. A classifier ensemble is generated with the aid of multiple-training learnings for the same assignment after which combining their predictions as verified later inside the study. There are many ways in which the ensembles may be generated with the ensuing output from every classifier then blended to classify new instances. In this way, we can evaluate both the dataset and the performance of the strategies used for prediction [8].

K-Fold cross-validation technique

In order to improve the accuracy, Cross-validation is a fitted technique to be used. It is a statistical process that can be utilized for estimating the aptitude of machine learning models. It constitutes a very popular method for applied machine learning when comparing and choosing a model for a given problem of prognostic modeling, because it is simple and can be easily implemented, yields the estimates of abilities that have a remittent bias unremarkably than other distinct approaches. In a resampling process that is used to analyze the machine learning approaches on a small record set, the cross-validation method is counted [9]. It is a commonly used technique because it is straightforward in understanding and because its outcomes can usually be less biased than those obtained from other alternative approaches. It constitutes a very straightforward "train"/"test" split.

PRINCIPAL COMPONENT ANALYSIS

The Principal Component Analysis (expansion of Karhunen-Loeve) is an appropriate related face recognition approach, since it distinguishes the variability between human faces, which may not be instantly apparent. PCA does not undertake to categorize faces by means of familiar geometrical differentiation, such as the width of an eyebrow or the length of the nose. Instead, the use of PCA has been studied by a tough and short face to see that 'variables' account for the variation of faces [10]. The value of describing the variance-covariance composition through certain linear combinations of the originally constructed variables is understood to be PCA. Due to the prominent power, PCA is capable of reducing and interpreting data [10]. For example, if an area contains a face area of 100×100 pixels, it will turn out to be perfectly defined by merely 40 eigen values. In each image, each individual value reveals the importance of each Eigenface. In addition, the entire perception (i.e. recognition) process can currently be accomplished by using only 40 eigenvalues to represent a face rather than manipulating the 10000 values found in a 100×100 image. Figure 1 demonstrates how the face recognition using the Principal Component Analysis approach works.

Linear Discriminant Analysis

Both PCA and ICA use many facial class figures. An important approach is used by Linear Discriminant Analysis to explain the-face-vector-spaces (Fig. 2). This differentiates people's names, but recognizes the faces of the same people [11, 12].
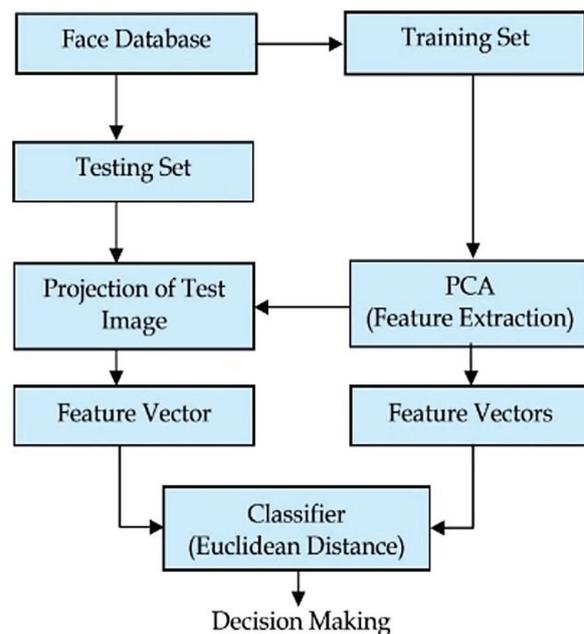

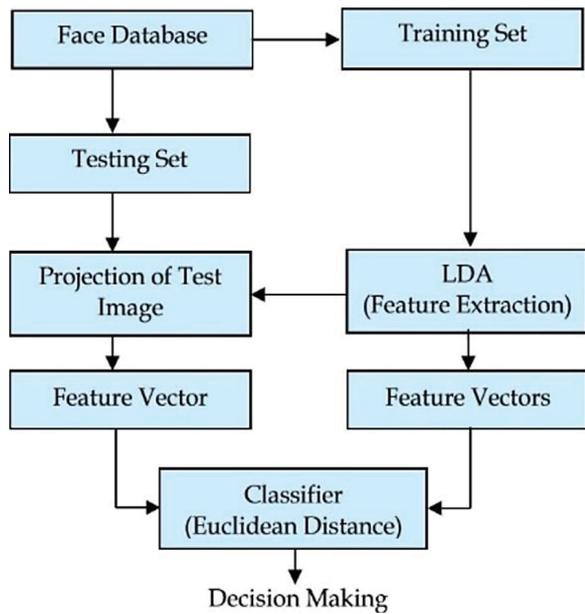**Fig. 1.** Face recognition using PCA approach

Fig. 2. Face recognition using LDA approach

Support Vector Machine

In 1992, Boser, Guyon, and Vapnik introduced the first Support Vector Machine to COLT-922. SVM is an approach that is widely used in machine learning. A very useful technique that is used in both classification problems and for regression problems [13÷15]. The wide use of SVM in classification issues shown in Figure 3. In our case, each data object was plotted considering a point in the n-dimensional region (where n is considered to be the large range of its characteristics you have) being the value of each characteristic, it is important to be the value of an individual coordinate. Afterwards, by seeking the hyper-plane that separates the 2 classes, we
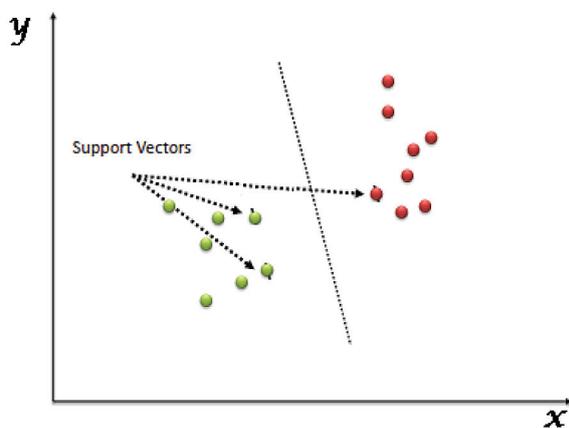
begin to carry out classification. Support Vector machines are simply individual observation coordinates. Most commonly, Support vectors are the frontier that does the segregation of the 2 classes very well (hyperplane/line).

Nearest Neighbor Algorithm

The 1-nearest neighbor algorithm in pattern recognition is a non-parametric approach used for regression and classification. The input consists of the nearest k training examples within the feature-space in each case. If 1-NN is used for classification or regression, the output depends on it [16]. The output is a class membership while using k-NN classification. An item is categorized by a plurality vote of its nearest neighbor, and alongside the object that is designated to the class most usual amongst its 1-nearest neighbors (where 1 is a positive integer, generally small). If k = 1, then the object is without a doubt assigned that alone nearest neighbor. In 1-NN regression, the term output is an attributed value for that specific object. This particular value is the average of the values of 1-nearest neighbor. 1-NN is a form of instance-based learning-method, or lazy learning method, where the function is uniquely approached localized and all the computations are differing till classification. Here, both classification and regression, constitute beneficial approaches that may be designate weights to the contributions of their neighbors; therefore, the contribution of nearest neighbor needed to provide more than averagely used. For instance, an average weighting scheme is composed of contributing each neighbor a weight of one/d, where d is called the distance to its neighbor.

Data set

This directory includes a collection of faces taken at the Olivetti Research Laboratory in Cambridge, UK, between April 1992 and April 1994 [19]. The data set is an open set. The directory also consists of 10 different images of 40 distinctive subjects. Some of the images have been taken at different time-frame for some specific subjects, varying lighting slightly, facial expressions (for example, in different poses with open or closed eyes, smiling and non-smiling faces), and facial details (wearing glasses and without glasses). All the images are taken on a dark homogeneous background, and the subjects are in

Fig. 3. SVM class dividing boundry

an upright place, frontal position with tolerance for some side movement [9, 17]. These images can be viewed with the 'xv' program which is very handy and the format of all these images is in the PGM, each image-file is 92x112 size, 8-bit grey levels, and all these images are structured in 40 directories (one for each subject) named as sX where X shows the subject number (between 1 and 40). In each directory there are 10 different images of the selected subject.

Purpose system

In the diagram (Fig. 4) 5 cases are shown to present our comparison. The common feature for all is that the dataset has been fed to K-fold algorithm to shuffle the images and increase its complexity. K-fold here shuffles the images and then split them in 5 different batches. Here, a single batch becomes a single fold, which is an input to a selected algorithm. Each fold has a set of training data and testing data, against which an accuracy value of algorithm is calculated. If there are 5 folds, this means the algorithm will receive input data 5 times from 5 different batches (folds) and accuracy is compounded. In the end, an average accuracy is calculated by summing 5 values of accuracy and dividing them by 5.

The first one is PCA+1NN. The train and test data have been fed to the PCA. PCA calculates the 70 strongest features (columns), which has

highest correlation between them. These features go to the nearest neighbor algorithm with a fixed parameter of finding 1st nearest neighbor. The 1NN algorithm first decides the label for the images by checking how close an image is to a certain label. Then, it compares its prediction with the original label to formulate accuracy. This process is performed 5 times, as the folds are set to be 5 and an average accuracy is calculated at the end of execution.

The second one on the list is LDA+1NN. LDA has been fed with training and testing data. Similar to PCA, it also fetches 70 strongest features among the data, but in contrast to PCA, LDA attempts to find a feature subspace that maximizes class separability. These features go to nearest neighbor algorithm with a fixed parameter of finding 1st nearest neighbor. The 1NN algorithm first decides the label for the images by checking how close an image is to a certain label. Then, it compares its prediction with the original label to formulate accuracy. This process is carried out 5 times, as the folds are set to be 5 and an average accuracy is calculated at the end of execution.

PCA+LDA+1NN, first data has been given to PCA to obtain 70 strongest features, and then those processed features are fed to LDA. It fetches 25 strongest features among the 70 features from PCA. These 25 features go to nearest neighbor algorithm with a fixed parameter of finding 1st nearest neighbor. The 1NN algorithm first decides
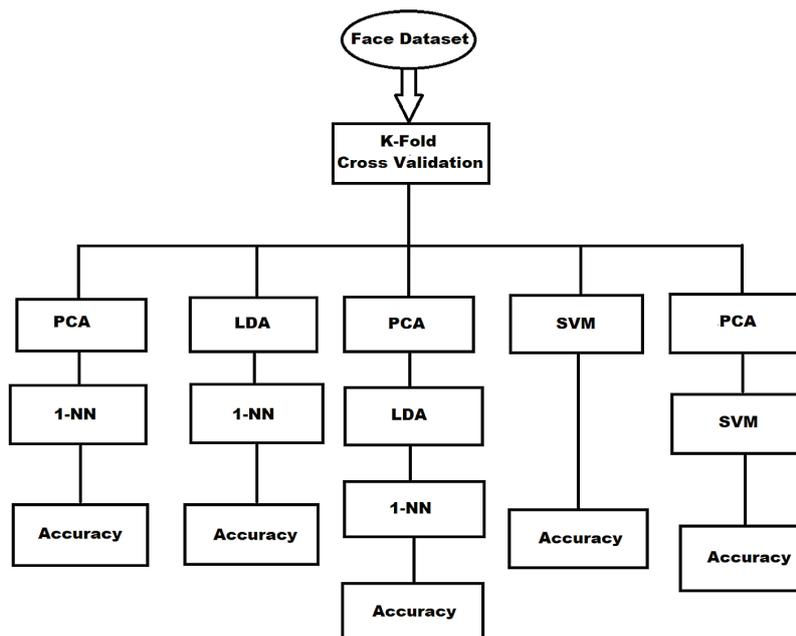

Fig. 4. Purpose system

the label for the images by checking how close an image is to a certain label. Then, it compares its prediction with the original label to formulate accuracy. This process is performed 5 times, as the folds are set to be 5 and an average accuracy is calculated at the end of execution.

Support Vector Machine has been popularly used for classification problems as a supervised learning algorithm. It takes training data to train, forms a hyper-plane to differentiate among classes and then maximizes on the margins of the plane from the values. After training, it predicts the testing data for classification. Accuracy is compounded for each fold and in the end we obtain an average accuracy from each fold.

Train and test data were fed to the PCA first. PCA calculates the 70 strongest features (columns), which have the highest correlation between them. Then, the output from PCA goes to SVM for training and classification. After the

SVM model has been trained, a test data set is given to SVM to predict the class of the images. The accurately classified images are then considered for accuracy from each fold. These compounded accuracies are then used to calculate the average accuracy.

Results

The results show that all the combinations of algorithms seem to perform very well on this dataset, but eventually LDA+PCA+1NN had better results (Fig. 5 and 6). If we perform comparison between PCA and LDA which are quite similar approach, they have individually given almost the same results (Fig. 7). This is due to same linear transformation mechanism, but LDA does maximum class separation by taking it to another axis. That causes it return a slight better result. Combining these two improves the score because
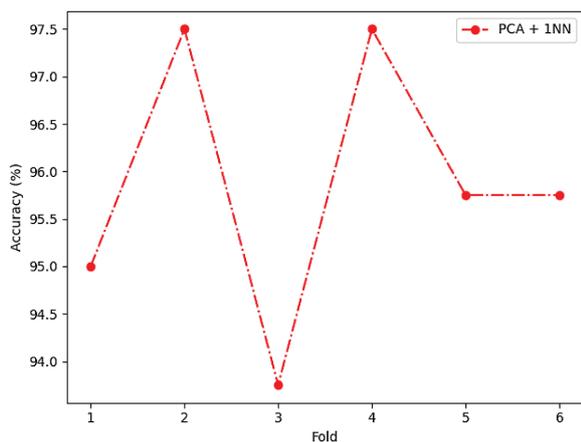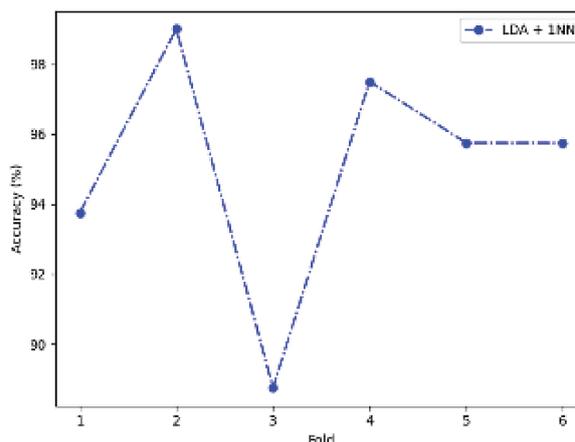

Fig. 5. PCA+1NN accuracy graph
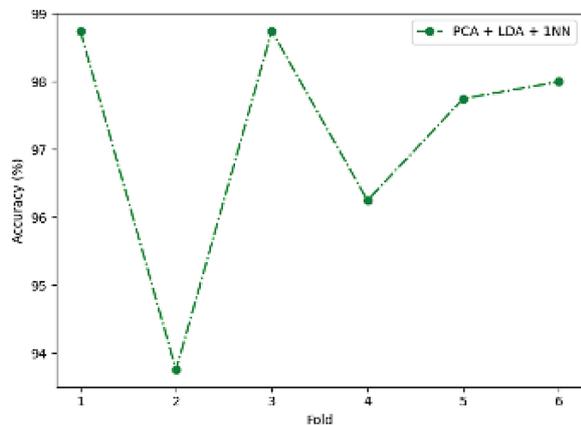

Fig. 6. LDA+1NN accuracy graph

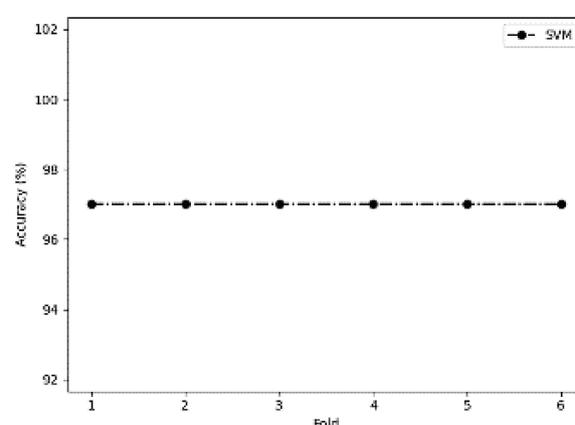**Fig. 7.** PCA+LDA+1NN accuracy graph
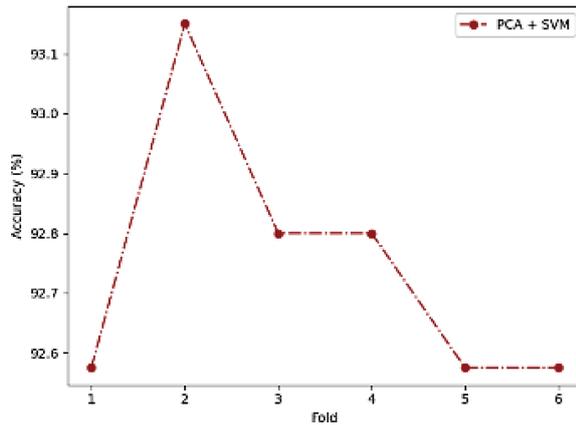

Fig. 8. SVM accuracy graph

Fig. 9. PCA+SVM accuracy graph
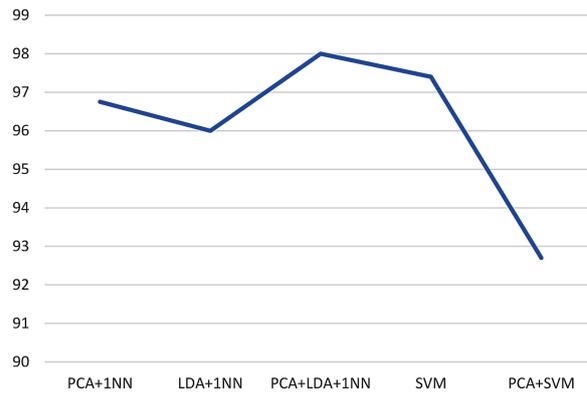

Fig. 10. Accuracy graph

Table 1. Average accuracy of different classifiers with 5-fold cross-validation

Classifiers	PCA+1NN	LDA+1NN	PCA+LDA+1NN	SVM	PCA+SVM
Average accuracy	96.75%	96%	98%	97.4%	92.7%

PCA obtains 70 most important features and those features are filtered out using LDA which take another 25 out of it. This makes a filtering effect, which makes better training and hence better results. SVM, on the other hand, uses the classification mechanism, separating by a hyperplane to distinguish two classes (Fig. 8). The results are very interesting, but considering it for small dataset it worked better. If the dataset was larger, it might not have performed in the similar fashion. Combining PCA and SVM would have shown better results, as they targeted features to train, but it lagged (Fig. 9). Some obvious reasons could be that the PCA features would not reconcile with original dataset features in terms of learning for SVM (Fig. 10). The average accuracy of different classifiers with 5-fold cross-validation was shown in Table 1.

CONCLUSION

This research has compared various image recognition algorithms on a given data set and the accuracy rate has been obtained on various folds. The data set contain images with certain variation like light or no light, glasses or without glasses or a photo with a smile or without a smile. The algorithm has achieved different accuracy rates on the given data set with respect to different folds. The accuracy rate has been compared and analyzed in the results section. It is clear from the graphs and

result tables that greater accuracy is obtained with the use of PCA, LDA and INN, as compared with all other used classifiers (98%).

REFERENCES

1. Buriro A., Crispo B., Conti M., A bimodal behavioral biometric-based user authentication scheme for smartphones. Journal of Information Security and Applications, 44, 2019, 89-103.

2. Gupta S., Gupta V.K., Lamba O.S., Digital image processing on face recognition using PCA. JETIR June 2019, 6(6).

3. Naik R., Pratap Singh D., Chaudhary J., A survey on comparative analysis of different ICA based face recognition technologies. Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2018, pp. 1913-1918, doi: 10.1109/ICECA.2018.8474860.

4. Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Tian-Hui Ma, Patch-based principal component analysis for face recognition. Computational Intelligence and Neuroscience, 2017, Article ID 5317850. https://doi.org/10.1155/2017/5317850.

5. Torkhani, G., Ladgham, A., Sakly, A. et al. A 3D–2D face recognition method based on extended Gabor wavelet combining curvature and edge detection. SIViP 11, 2017, 969–976. https://doi.org/10.1007/s11760-016-1046-7.

6. Benhida E.K., Boulahoual A., Face recognition of face images with hidden parts using Gabor wave-

lets F. and PCA", Periodicals of Engineering and Natural Sciences, 6(2), 2018.

7. Payal P., Goyani M.M., A comprehensive study on face recognition: methods and challenges. The Imaging Science Journal, 68(2), 2020, 114-127. DOI: 10.1080/13682199.2020.1738741.

8. Faridi M.S., Zia M.A., Javed Z., 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), 2021.

9. Kerbaa T.H., Mezache A., Oudira H., Model selection of sea clutter using cross validation method 9. Procedia. Computer Science, 158, 2019, 394-400. https://doi.org/10.1016/j.procs.2019.09.067.

10. Zhang Y., Xiao X., Yang L., Xiang Y., Zhong S., Secure and efficient outsourcing of PCA-based face recognition. IEEE Transactions on Information Forensics and Security, 15, 2020, 1683-1695. doi: 10.1109/TIFS.2019.2947872.

11. Barnouti N.H., Matti W.E., Al-Dabbagh S.S.M., Face detection and recognition using Viola-Jones with PCA-LDA and square Euclidean distance. International Journal of Advanced Computer Science and Applications, 7(5), 2016.

12. Borade S.N., Deshmukh R.R., Ramu S., Face recognition using fusion of PCA and LDA: Borda count approach. 24th Mediterranean Conference on Control and Automation (MED), Athens, 2016, 1164-1167, doi: 10.1109/MED.2016.7536065.

13. Thissen U., Pepers M., Üstün B., Melssen W.J., Buydens L.M.C., Comparing support vector machines to PLS for spectral regression applications. Chemometrics and Intelligent Laboratory Systems, 73(2), 2004, 169-179.

14. Zhihua Xie, Guodong Liu, Infrared face recognition based on binary particle swarm optimization and SVM-wrapper model. Proc. SPIE 9674, AOPC 2015: Optical and Optoelectronic Sensing and Imaging Technology, 96740J (15 October 2015). https://doi.org/10.1117/12.2197388

15. Dadi H.S., Pillutla G.K.M., Improved face recognition rate using HOG features and SVM Classifier. IOSR Journal of Electronics and Communication Engineering, 11(4), 2016, 34-44.

16. Dhinakaran, M.S., Thirumaran, J., Evaluation of profile-based personalized web search using KNN and ECC. International Journal of Applied Engineering Research, 13(19), 2018, 14411-14416.

17. Friedman, L. et al. Method to assess the temporal persistence of potential biometric features: Application to oculomotor, gait, face and brain structure databases. PLoS One, 12(6), 2017, e0178501.

18. Key Parameters and Configuration for Streaming Replication. https://www.enterprisedb.com/blog/high-availibility-parameters-configuration-streaming-replication-postgresql

19. GitHub.efflorez/facerecognition. https://github.com/efflorez/facerecognition

20. Ramezan, C.A, Warner T.A., Maxwell A.E., Evaluation of sampling and cross-validation tuning strategies for regional-scale machine learning classification. Remote Sensing 11(2), 2019, 185-206.

21. Axnick, K.B., Ng, K.C. 2005. Fast face recognition. Image and Vision Computing, 43-48.

22. Starovoitov V.V., Samal D.I., Briliuk D.V., Three approaches for face recognition. The 6th International Conference on Pattern Recognition and Image Analysis, October 21-26, 2002, Velikiy Novgorod, Russia, pp. 707-711.