British Journal of Applied Science & Technology
13(4): 1-10, 2016, Article no.BJAST.21960
ISSN: 2231-0843, NLM ID: 101664541

SCIENCEDOMAIN *international*
www.sciencedomain.org

Normalized Independent Component Analysis for Face Recognition

I. Oyedeji Ayo[1*], O. Omidiora Elijah[2] and Oladipo Oluwasegun[3]

[1]Department of Computer Engineering, Ogun State Institute of Technology, Igbesa, Ogun State, Nigeria.
[2]Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.
[3]Centers for Information Technology and Management, Yaba College of Technology, Yaba, Lagos State, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Authors IOA and OOE designed the study. All the authors worked on development of the NICA algorithm and the flow to ensure the system works. Authors IOA and OO wrote the first draft of the manuscript and managed the comparison analysis. Author OOE supervised the project. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJAST/2016/21960
Editor(s):
(1) Samir Kumar Bandyopadhyay, Department of Computer Science and Engineering, University of Calcutta, India.
Reviewers:
(1) Ayşsegul Uçar, Firat University, Turkey.
(2) Anonymous, University of Fez, Morocco.
(3) Wen-Yeau Chang, St. John's University, Taiwan.
Complete Peer review History: http://sciencedomain.org/review-history/12515

Original Research Article

Received 11th September 2015
Accepted 16th November 2015
Published 2nd December 2015

ABSTRACT

Aims: To design a Face Recognition System (FRS) using combination Independent Component Analysis (ICA) and Artificial Neural Network - Normalized ICA. In order to improve the performance of a conventional ICA which suffers the drawback of ranking the energies of the generated features.
Study Design: The FRS was simulated using Matlab 2011 version. An algorithm was developed which combines the ability of the conventional ICA with ANN to generate final predictions. The ANN serves as a region finder and generated likely predictions associated with face image classes. Hence, reduced search space during testing.
Place and Duration of Study: Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria, between June 2014 and July 2015.

**Corresponding author: E-mail: ayooyee@yahoo.com;*

Methodology: 40 individuals face images were captured in an uncontrolled environment. The face database comprises of 10 images for each subject taken at different times. The images were pre processed by cropping to different sizes (92*92; 92*100; 92*112 pixels respectively) and removing unwanted background. During testing Euclidean distance was used as similarity measure and faces were classified as "known" if less or equal to the threshold value set else "unknown".
Results: The recognition accuracies at dimension 92*92 are 86.00% and 95.00% for ICA and NICA-based system using 30 principal components, 86.50% and 96.00% using 60 principal components at the same dimension respectively. At dimension 92*112 a recognition accuracy of 90.00% and 98.00% was obtain for ICA and NICA-based system, 91.00% and 98.00% using 60 principal components at the same dimension respectively. At cropped dimension 92*92 it took an average of 0.0096s and 0.0095s using 30 principal components to recognize a test image in ICA- and NICA-based, 0.0086s and 0.0085s using 60 principal components at the same dimension respectively and at cropped dimension of 92*112 it took an average of 0.0102s and 0.0098s using 30 principal components, 0.0106s and 0.0099s using 60 principal components at the same dimension respectively
Conclusion: The developed NICA-based system has better recognition accuracy than a conventional ICA-based system and also recognizes face images faster.

Keywords: Normalized; independent component analysis; face recognition.

1. INTRODUCTION

In recent times, the need of biometric security system is heightened for providing safety and security against terrorist attacks, robbery, etc. The demand of biometric system has risen due to its strength, efficiency and easy availability. One of the most effective, highly authenticated and easily adaptable biometric security systems is facial feature recognition [1].

The recent advance in computer vision, pattern recognition and image processing, face recognition is one of the most popular research topics. The reason behind this is that among the various biometric security systems based on finger print, iris, voice or speech, signature, etc., face recognition seems to be the most universal, non-intrusive, and accessible system [2]. It is easy to use, can be used efficiently for mass scanning which is quite difficult in case of other biometrics, and also increases user-friendliness in human-computer interaction.

The earliest works on this subject were made in the 1950's in psychology. There are other issues like face expression, interpretation of emotion or perception of gestures. In the past 20 years, significant advances have been made in design of successful classifier for face recognition using template matching algorithm [3]. However the diversity of the face image patterns makes it difficult to create robust recognition systems and the complexity of the algorithms makes them tedious to implement [4]. A successful face recognition methodology depends heavily on the particular choice of the features used by the (pattern) classifier. Feature selection in pattern recognition involves the derivation of salient features from the raw input data in order to reduce the amount of data used for classification and simultaneously provide enhanced discriminatory power. This research work developed a Normalized Independent Component analysis in image processing, with particular reference to human faces and then compares its results with the convectional Independent Component Analysis in terms of recognition accuracy and recognition time.

Independent Component Analysis (ICA) has emerged recently as one powerful solution to the problem of blind source separation i.e. different features classification. ICA is an unsupervised feature extraction technique whose goal is to find which components are statistically independent from each other as possible. ICA is the generalization of Principal Component Analysis (PCA). In high-dimensional applications (set of features), the ICA pipeline actually contains PCA process (for dimension reduction), whitening process (for scaling), and pure ICA process. PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. The principal components are orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric. The implementation of PCA in face recognition is commonly called Eigenface.

In a task such as face recognition, much of the important information may be contained in the high-order relationships among the image pixels. PCA is only sensitive to the second order statistics by de-correlating the input data using second-order statistics and thereby generates compressed data (dimension reduction) with minimum mean-squared reprojection error using Eigen faces [5] for representation. ICA minimizes both second-order and higher-order dependencies in the input.

The Normalized Independent Component Analysis (NICA) uses the K-means Nearest Neighbor (KNN)) method for finding similar train instances to each test sample which helps in determining the energies of the different independent components and then computes some measures to determine the Power Recognition Factor (PRF) whose results are passed to the Artificial Neural Network (ANN) system for training. ANN is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because the neural network learns based on that input and output.

The KNN, PRF and normalized PRF were deployed to combat the problem of the convectional ICA. The problem includes the inability to determine the order and variance (energies) of the independent components. Ultimately, ANN learnt and predicted region for best recognition from the normalized PRF.

Facial Biometrics has been implemented using various approaches such has local feature analysis, Independent Component Analysis, and Principal Component Analysis and Linear Discriminant Analysis (LDA) but the algorithms can be further enhanced to provide reliable face recognition, maximize recognition rate; regardless of the facial orientation, face expression, illumination variation and face dressing. This research work presents the development of a Normalized -Independent Component Analysis algorithm and investigates its performance viz a viz a conventional ICA-based system with respect to recognition time and recognition accuracy.

2. LITERATURE REVIEW

The growing use of the Internet by individuals and organizations has presented opportunities for identity fraud, organized crime, money laundering, theft of intellectual property and a myriad of cybercrimes. The world has also witnessed an increase in bio-security incidents, border control incidents and terrorism. With this background, the ability to positively identify people is becoming more and more important. One key tool in this area is the use of biometrics. Humans have always identified each other by recognizing faces, voices or some other physical characteristic. Personal recognition or identification by a witness is also entrenched in our law and commercial structures. Now the use of biometric technologies is providing a means to positively identify or authenticate large numbers of people without having to primarily rely on human to human identification [6]. Whatever biometric technology, or combination of technologies are used, there are common basic procedures: this include Capture, extract, create template and compare. Capture is also known as the enrolment or registration process. In this process the individual provides a sample of the appropriate biometric, for example provide a fingerprint, repeats a phrase for a voice recognition system or provide an image for a face recognition system. This sample is analysed and key features extracted to create a template. In some cases, several samples are provided and the key features aggregated into the template, sometimes also known as a reference template. During Extraction, Ideally the sensors will extract sufficient data to ensure the key characteristics do not vary over time or with changes in the environment (such as a poorly presented fingerprint or poor lighting for facial recognition sensors). Devices typically take multiple samples of the presented biometric and average the results to produce the template. To minimize the size of the template and storage requirements and to speed searches, the sensors usually extract only features that tend to be unique. Ideally some form of "liveness" test should also be incorporated to minimize or defeat attempts to circumvent or compromise the system.

The template is the data representing the captured biometric and the reference against which later samples are checked. Templates are only data representing key or distinctive features of a biometric and are not a complete image or record of the original biometric (such as a fingerprint, voice recording or digital image). One analogy used is that biometrics is the body's passwords. They are usually small, which facilitates processing and speeds response times when referenced. The ease of enrolment and quality of the template are often the determining factors in the successful implementation and use of a biometric system. The comparison of the sample to templates and its ability to successfully discriminate between samples and templates is, perhaps, the most important part of a biometric system. Systems for matching have been used in a wide variety of applications for some time. Sometimes described as pattern matching, it can use a variety of techniques such as minimal distance, probabilistic analysis and neural networks. The comparison process can also add the current sample to the template, thus averaging results over a long period of use.

3. METHODOLOGY

This study presents the development of a Face Recognition System (FRS) by combining; ICA and ANN. The ICA uses the Principal component Analysis (PCA) for dimension reduction by considering the eigenvector associated with higher Eigen values in features extraction and the Artificial Neural Network (ANN) algorithm remove unimportant features in the eigen face space. This ensures that only relevant features are passed to the testing stage. The developed system used the Euclidean distance as similarity measure to determine if the face is recognized or not. The model diagram for the FRS developed is shown in Fig. 1.

The face database used in this simulation consists of 400 images. Ten (10) different images of 40 distinct individuals were taken and some subjects' images were taken at different times. The sequence of the steps involved in developing the system is described in the flow charts as shown in Figs. 2 and 3 respectively.

The training stage is preceded by series of pre-processing stage which includes: image cropping to remove background and unwanted parts of the image, image grayscale conversion and conversion of image data to vector form. Independent component analysis uses the Principal component analysis on the vector form to ensure that the face image dataset is represented by a reduced number of effective features and yet retains the most intrinsic information of the image data. The face database used has 200 images for training and each image has minimum resolution of 92*92(8464) and maximum of 112*92 (10304) pixel with 30 and 60 selected principal components after background removal. Sample images are shown in Fig. 4. It should be noted that only first 30 and 60 principal components were selected for the simulation. Independent component analysis was then applied on the eigen values to make the feature vector independent of another and to be able to handle the dependency that exists among feature which slow down classification process and reduce the recognition accuracy. The Independent features are then saved for further processing. During this study it was seen that ICA lacks the capacity to know the order (i.e. energies) of each independent component, this was also stated in [7]. In a bid to overcome this shortcoming, the already trained and saved Independent Components was then given back to the Independent Component Analysis (ICA) as

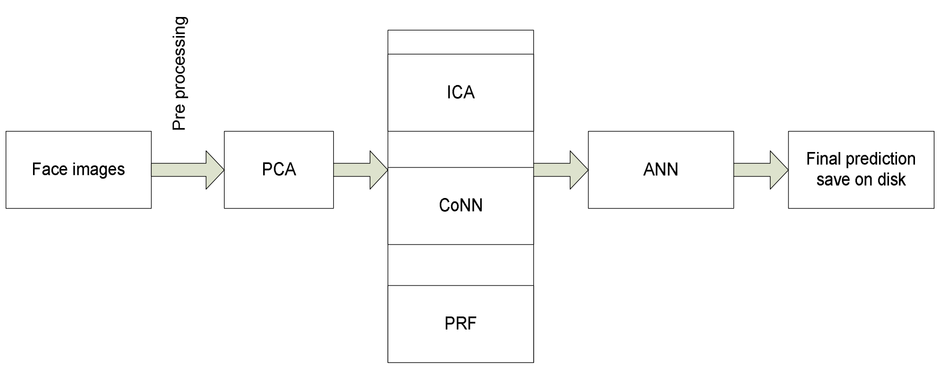


Fig. 1. The developed FRS model

CoNN = Computation of Nearest Neighbor; PRF = Power Recognition Factor test faces to obtain 5-Nearest Neighbors (NN) for all the 200 training datasets. Six (6) different measures were applied to the NN feature set generated from the previous stage to compute the Power Recognition Factor (PRF); this was carried out to train ICAs in different classification regions.

The measures include P1 = Number of correct recognized results in the first K results, P2 = the first incorrect result index, P3 = Distance between the first correct and the first incorrect result, P4 = Distance between the first incorrect result and its previous, P5 = Number of times that the instance is placed in a positive situations and N6 =Number of times that the instance is placed in a negative situations.

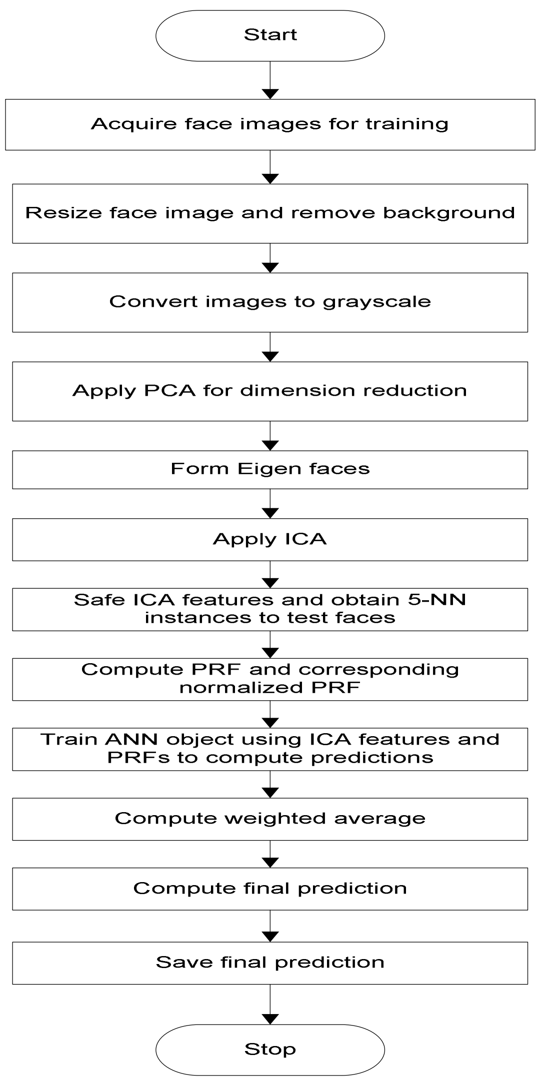


Fig. 2. Training stage flow chat

The PRFs were normalized to form nP1-nP5 and nN6 respectively. This serves as the targeted output of artificial neural network. ANN objects were trained using Independent component Analysis features as the input and the computed normalized power recognition Factor (PRF) as targeted outputs. ICA features were then simulated with created ANN objects to obtain predictions for each of the training instances. The weighted average of each training instances were also computed. The product of the predictions and the weighted average calculated formed the final predictions which were saved as NICA face space to be used during testing. In this, the ANN is a region finder. It created predictions from the normalized PRFs to form possible instances presented during testing. This reduced testing search space and ensure that only relevant instances are searched to know if the face image was recognized or not.

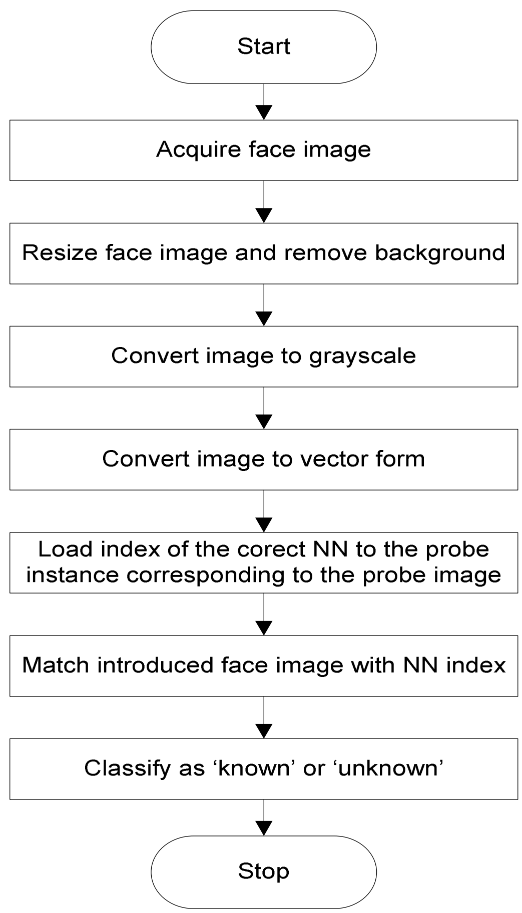


Fig. 3. Testing stage flow chat

During testing, images were projected into NICA face space after it has been pre-processed and the task of determining whether the image was recognized or not was carried out. This was achieved by performing distance measure using Euclidean distance between the projected test image and the index of the correct nearest neighbor (NN) to the instance corresponding to the probe or test image in the created NICA face space. The image that has the minimum Euclidean distance to the test image is selected. In this case the image was classified as "known" and if the image that aren't in the same class with test image was selected as having the minimum Euclidean distance to the test image, the image is said to be "unknown".

All the 400 images acquired were used in this research work. Five (5) images from each class of 10 are chosen for training while the other five were used in testing the system. The test images were introduced one at a time specifying the file name, to see if it will be classified as being "known" or "unknown" to the FRS. The performance metrics used are Average recognition accuracy, average recognition time and total training time. These were investigated as it varies with cropped image dimension to deduce optimum dimension for the face recognition system and the results were evaluated with the performance of a convectional ICA based recognition system.

The simulation process preceded in two stages namely the training stage and testing stage. During the training stage, the system used five (5) of the 10 images samples to form NICA feature space. This was done by iteration through the faces in the f1 to f40 folder of the project directory and training the files names 1jpeg, 2jpeg, 3jpeg, 4jpeg and 5jpeg respectively. The testing image sample were arranged thus; 5pjeg, 6jpeg, 7jpeg, 8jpeg, 9jpeg and 10jpeg were the five remaining images of the first subject, 15jpeg, 16jpeg, 17jpeg, 18jpeg, 19jpeg and 20jpeg as the remaining sample of the second subject respectively. These were all stored and arranged in the project directory of the system's local drive local drive.

4. RESULTS AND DISCUSSION

The dimension of the cropped image influences the performance of both NICA and ICA-based FRS. Table 1 shows the summary of the effect of increase in the cropped dimension from 92*92 through 92*112 for both NICA and ICA-based systems. The recognition accuracies at dimension 92*92 are 86.00% and 95.00% for ICA and NICA-based system using 30 principal components, 86.50% and 96.00% using 60 principal components at the same dimension respectively. At dimension 92*112 a recognition accuracy of 90.00% and 98.00% was obtain for ICA and NICA-based system, 91.00% and 98.00% using 60 principal components at the same dimension respectively. This implies that recognition accuracy of both systems increases with increase in crop dimension. This is because more features will be available to identify the face image with an increase in cropped image dimension. The NICA-based system was found to have better recognition accuracy; this was due to the fact that the only relevant features were made available for recognition in the NICA feature set.

The time of recognition of images introduced to the system was well captured, as the program evaluated the difference between the time image was recognized and the time the search commences. The system was made to iterate through images 5pjeg, 6jpeg, 7jpeg, 8jpeg, 9jpeg and 10jpeg through 396jpeg, 397jpeg, 398jpeg, 399jpeg and 400jpeg in the project directory; which were the equivalent testing images. The time to recognize each image was captured and cumulated to calculate the average time it takes to recognize a test face image. At cropped dimension 92*92 it took an average of 0.0096s and 0.0095s using 30 principal components to recognize a test image in ICA-and NICA-based, 0.0086s and 0.0085s using 60 principal components at the same dimension respectively and at cropped dimension of 92*112 it took an average of 0.0102s and 0.0098s using 30 principal components, 0.0106s and 0.0099s using 60 principal components at the same dimension respectively as shown in Table 2. This shows that NICA-based system optimizes search time better than the ICA counterpart. This is due to the fact that the features presented for NICA search had only relevant features to search through. The time incurred while training the NICA-based system was captured as the difference between the cpu-time at the commencement of the first training and cpu-time at the conclusion of the last training image.

At dimension 92*92 a total time of 20.7481s and 26.3174s using 30 principal components was incurred while training the ICA-and NICA-based system, 41.5743s and 47.1435s using 60 principal components was incurred at the same dimension respectively while at dimension 92*112 a total time of 22.1677s and 27.9398s using 30 principal components was incurred while training the NICA- and ICA- based system, 45.2403s and 49.1091s using 60 principal components was incurred at the same dimension respectively as shown in Table 3. This indicated that the ICA- based system trains faster than the NICA counterpart. This could be explained by the time incurred in ANN training and normalization process in removing irrelevant features in the NICA feature space.

Table 1. Recognition accuracy for both NICA-and ICA-based system

Image dimensions	PC	NICA recognition accuracy (N)	ICA recognition accuracy (I)	N - I	Percentage increment N - I * (100 / I)
92 * 92	30	95.00	86.00	9.00	10.46
	60	96.00	86.50	9.50	10.98
92 * 100	30	97.00	87.00	10.0	11.49
	60	97.50	88.50	9.00	10.16
92 * 112	30	98.00	90.00	8.00	8.88
	60	98.00	91.00	7.00	7.69

** Pc = Principal components*

Table 2. Average recognition time for both NICA-and ICA-based system

Image dimensions	PC	NICA recognition time (N_t)	ICA recognition time (I_t)	N_t - I_t	Percentage increment N_t - I_t * (100 / I_t)
92 * 92	30	0.0095	0.0096	0.0001	10.46
	60	0.0085	0.0086	0.0001	10.98
92 * 100	30	0.0093	0.0095	0.0002	11.49
	60	0.0089	0.0099	0.001	10.16
92 * 112	30	0.0098	0.0102	0.0004	8.88
	60	0.0099	0.0106	0.0007	7.69

Fig. 4. Sample images after background removal

Table 3. Training Time for both NICA-and ICA-based system

Image dimensions	PC	NICA training time (N_T)	ICA training time (I_T)	$N_T - I_T$	Percentage increment $N_T - I_T * (100 / I_T)$
92 * 92	30	26.3174	20.7481	5.5693	26.842
	60	47.1435	41.5743	5.5692	13.395
92 * 100	30	27.9398	22.1677	5.7721	26.038
	60	49.1091	45.2403	3.8688	8.551
92 * 112	30	29.7026	23.2129	6.4897	27.957
	60	55.8172	50.6691	5.1481	10.160

5. CONCLUSION

In conclusion the simulation reveals that the NICA- based system developed had a better recognition accuracy and better recognition time than the conventional ICA- based system but ICA- based system trains faster than the NICA-based system. Generally recognition accuracies increases with increase cropped face image dimension used.

CONSENT

It is not applicable.

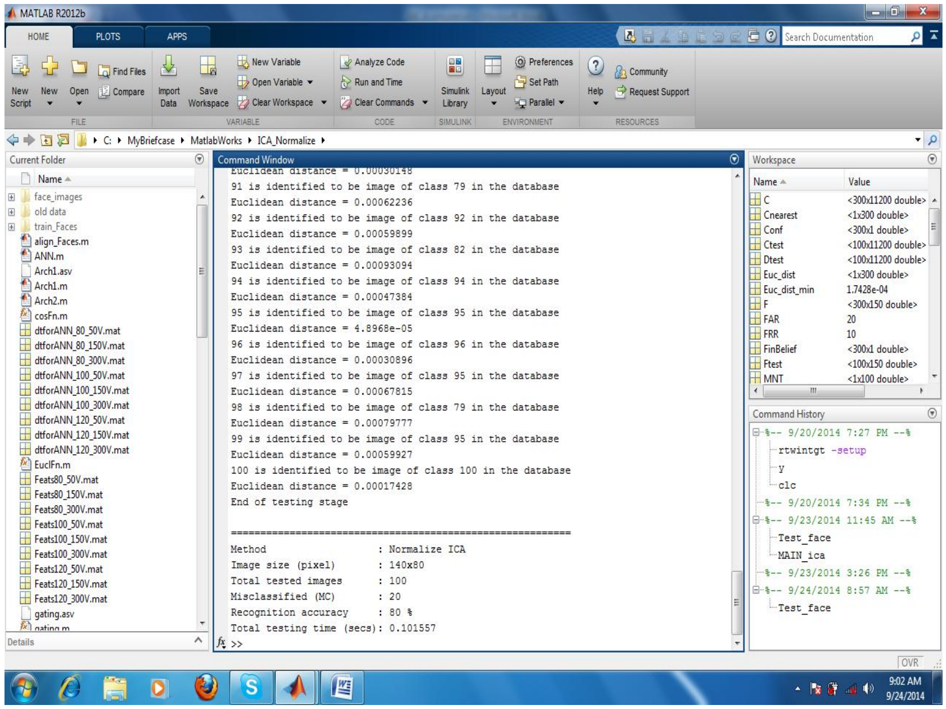
COMPETING INTERESTS

Authors have declared that no competing interests exist.

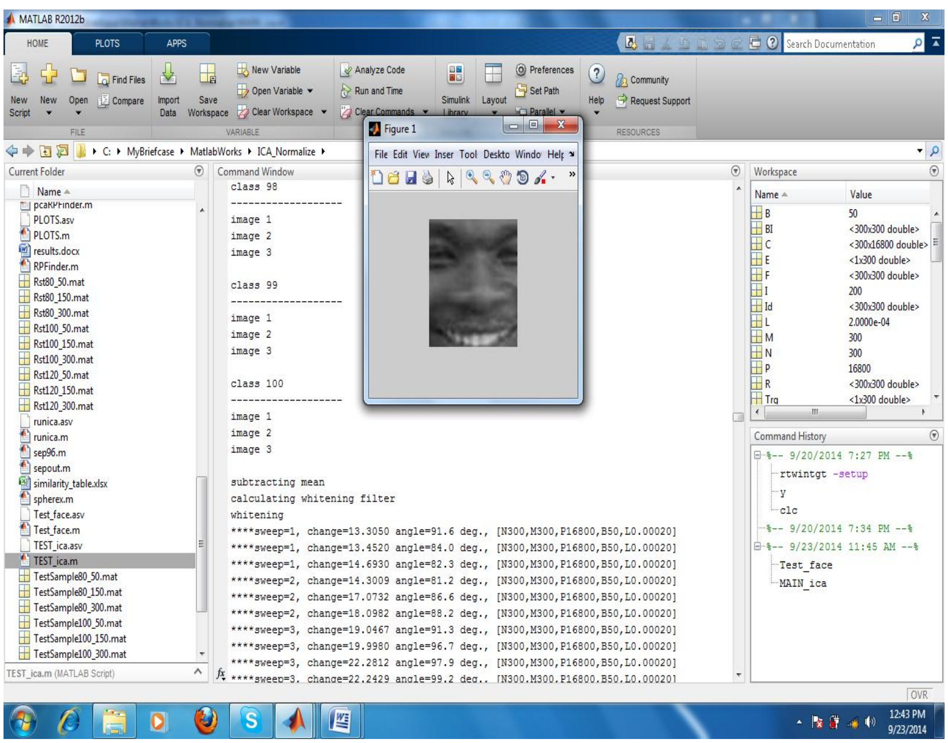
REFERENCES

1. Omidiora EO. A prototype of knowledge-based system for black face recognition system using principal component analysis and fisher disriminant algorithms. PhD. Computer Science Thesis, Ladoke Akintola University of Technology, Ogbomoso, Nigeria; 2006 (In press).
2. Oladipo O. Development of a GPCA-based recognition system for face recognition of eigenfaces. M.Tech Computer Science Thesis. LadokeAkintola University of Technology. Ogbomoso, Nigeria; 2014 (In press)
3. Aria VN, Monson H. Hidden Markov models for face recognition. IEEE International Conference on Acoustic, Speech and Signal Processing. 1998;5: 27224-2721.
4. Omidiora EO, Fakolujo OA, Ayeni RO, Olabiyisi SO, Arulogun OT. Quantitative evaluation of principal component analysis and fisher discriminant analysis techniques in face image. Journal of Computer and its Applications. 2008;15(1):37-22.
5. Turk M, Pentland A. Eigenfaces for recognition. J. Cognitive Neurosci. 1991; 3(1):86–71.
6. Birgit K. Biometric authentication systems today and in the future. Computer Science Department University of Auckland; 2005.
7. Mehran N, Mansour JA. Solution of combining several classifiers for face recognition Proceedings of the World Congress on Engineering and Computer Science WCECS; October 24-26, San Francisco, USA; 2007.

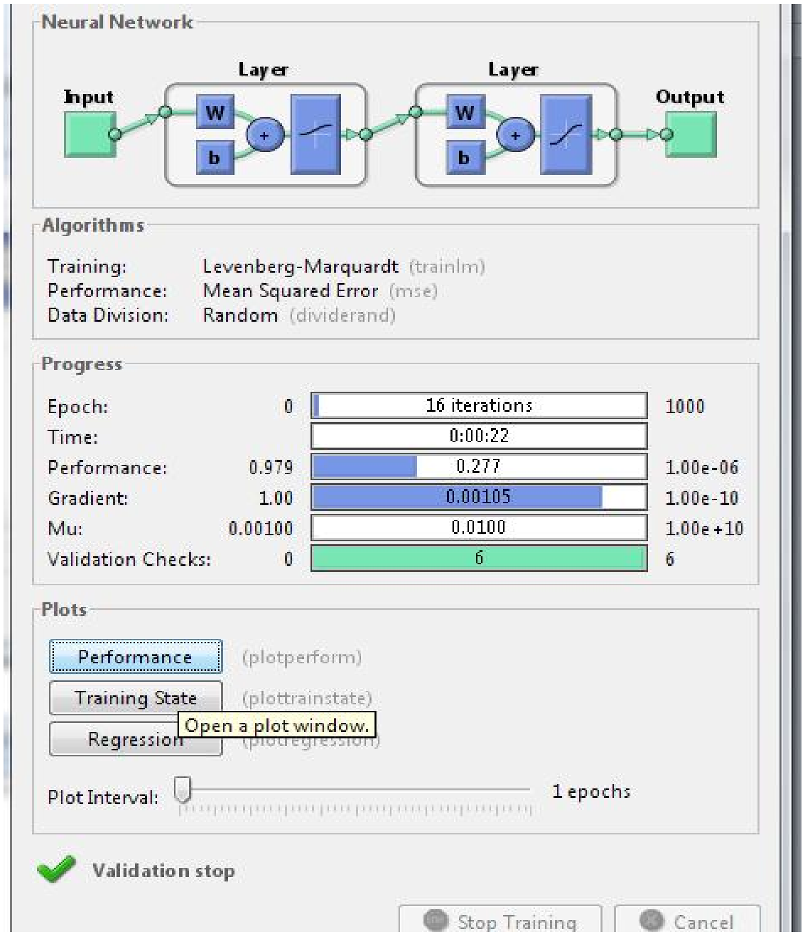
APPENDIX



Batch testing after training NICA-based FRS



Individual testing after storing NICA features on local disc



Snapshot during ANN simulation

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/12515