



A Novel Facial Image Deviation Estimation and Image Selection Algorithm (Fide-Isa) for Training Data Reduction in Facial Recognition System

Hassan Opotu Siyaka, Olumide Owolabi and Bisallah, I. Hashim
Department of Computer Science,
University of Abuja, Nigeria

ABSTRACT

Deep learning models have been successful and shown to perform better in terms of accuracy and efficiency for facial recognition applications. However, they require a huge amount of data samples that were well annotated to be successful. Their data requirements have led to some complications which include increased processing demands of the systems where such systems were to be deployed. Furthermore, the unavailability of well-annotated data has also affected the success of deep learning facial recognition models. Reducing the training sample sizes of deep learning models is still an open problem. This paper proposes in reducing the number of samples required by the convolutional neural network used in training a facial recognition system. A Facial Image Deviation Estimation and Image Selection Algorithm (FIDE-ISA) was developed to select appropriate training samples incrementally based on their facial deviation. The model was tested using an online dataset called Face94 facial image dataset. The results indicated that training a CNN model with at least 3 images per individual produces a 100% accuracy of recognition. This is an indication of the effectiveness of the developed algorithm.

ARTICLE INFO

Article History

Received: August, 2020

Received in revised form: January, 2021

Accepted: February, 2021

Published online: May, 2021

KEYWORDS

Facial Recognition, Convolutional Neural Network, Deep Learning, Facial Image Deviation Estimation

INTRODUCTION

Face recognition is a biometric method used in recognizing an element using facial features. It is characterized by the high accuracy of a physiological approach without being intrusive (Ogbuju et al., 2020). Face recognition as a biometric method has brought the attention of many researchers in different faces ranging from computer vision, psychology, business, security and image processing (Alqadi et al., 2020; Coskun et al., 2017). Despite their importance, some factors affect the performance of face recognition systems. This includes changes in postures, scene illumination, orientation, the field of view and database sizes (Ogbuju et al., 2020; Almabdy & Elrefaei, 2019; Arafah et al., 2020; Coskun et al., 2017).

Several techniques have been proposed for face recognition applications, this includes deep learning which has shown better performance than other models in terms of

accuracy and processing (Coskun et al., 2017; Filippidou & Papakostas, 2020; Ling et al., 2020; Marsot et al., 2020; Swapna et al., 2020; Tao et al., 2019). CNN's had been applied successfully to solve several computer vision problems with great success attributed to using large-scale data samples (Pei et al., 2019). However, due to difficulties in sample collection and the requirement for training with large-sample data, deploying CNNs for daily attendance management remains a challenge (Pei et al., 2019; Parkhi et al., 2015). Training a deep learning model with a huge amount of data allows the model to learn from different variations that exist in the images and that may be required for recognition (Peng et al., 2016).

To address the sample requirement issues of CNN face recognition, Pei et al., (2019) proposes the application of data augmentation to enlarge the datasets by deploying several geometric transformations, filters and

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

enhancement mechanisms. The result indicated an improvement in the accuracy of the system, however, the complexities of the system also increased. Reducing the computational complexities of deep learning networks, and their dependence on large accurately annotated databases for training deep learning models is still an open research area that requires urgent attention (Nagpal et al., 2019). Peng et al., (2016) applied modelling methods to reduce the need for huge datasets. This was achieved in three phases; application of spatial transformation learning for aligning faces spontaneously and design of a robust face recognition system. This paper proposes the reduction of the number of training samples a CNNs will require for training and still obtain a state of the art performance. A Facial Image Deviation Estimation and Image Selection Algorithm (FIDE-ISA) was developed to select appropriate training samples based on their deviation in an incremental manner. The method adopted in this paper will ensure that images with closely related features, that is, minimum

deviation are avoided while those with wider deviations are selected to capture the diversity of the samples. The remaining part of the paper is organized as: In section two, methods and materials used were described, and in section three, the results and discussion were presented.

MATERIALS AND METHODS

Facial Image Deviation Estimation and Image Selection Algorithm (FIDE-ISA)

FIDE-ISA is an algorithm that was designed to estimate the deviations within facial images of an individual and selects the best images for training a deep learning facial recognition model. To achieve this successfully, five facial landmarks were detected which are:

- i. Left eye
- ii. Right eye
- iii. Nose
- iv. The left side of the mouth
- v. The right side of the mouth

Figure 1 shows a facial image with the facial landmarks of interest.

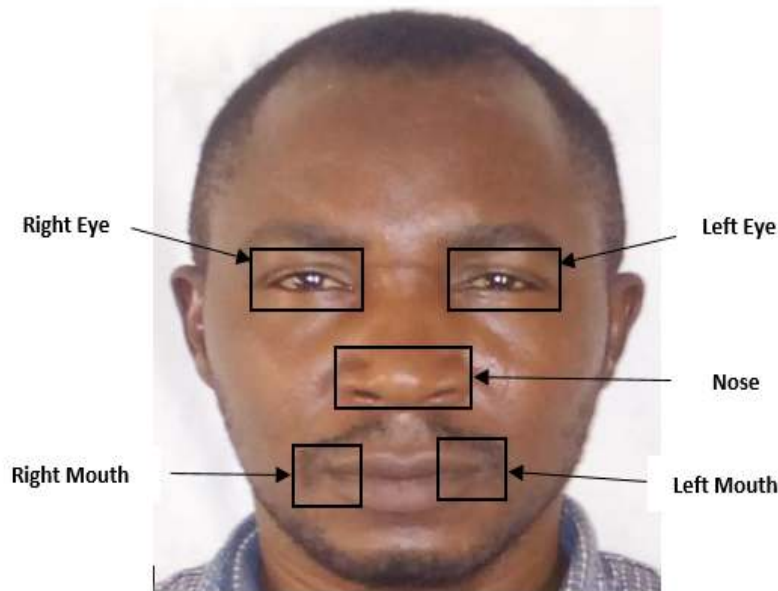


Figure 1: Facial landmarks

The facial landmarks used in Figure 1 were detected using the MTCNN model which stands for (Multi-task Cascaded Convolutional Neural Network) used in face detection. MTCNN

was adopted for both facial landmark detection and face detection because it produces better results than any face detection model. After detecting the facial landmarks, the Euclidean

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

distances between the coordinates of the base image and the deviated image is calculated as illustrated in Figure 2.

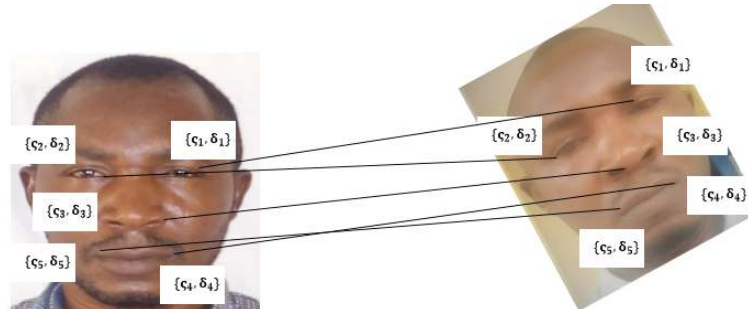


Figure 2: Coordinates of Facial Landmarks for Base Image and Deviated Image

Given a facial image 'I', with five facial landmarks, the landmarks can be represented as:

$$L_I = [\{c_1, d_1\}, \{c_2, d_2\}, \{c_3, d_3\}, \{c_4, d_4\}, \{c_5, d_5\}] \quad (1)$$

Where $c_1: c_5$ are the x-coordinate of the left eye, right eye, nose, left mouth and right mouth respectively. Similarly, $d_1: d_5$ are the y-coordinates of the left eye, right eye, nose, left mouth and right mouth respectively. The landmarks of the image I can be represented as a vector V as follows:

$$V_i = \{c_1, d_1\}, \{c_2, d_2\}, \dots, \{c_N, d_N\} \quad (2)$$

Where N = Number of facial landmarks, $i=1:N$. For M number of images, j^{th} image landmarks can therefore be represented as:

$$V_j = \{V_1\}, \{V_2\}, \{V_3\}, \dots, \{V_M\} \quad (3)$$

The Euclidean distance (D) between two images, the base image (I_b) and the deviated image (I_d) was calculated using Equation 3.4.

$$D_{bd} = \sqrt{\sum_{i=1}^N (V_b - V_d)^2} \quad (4)$$

For (M - 1) deviated images, the Euclidean distances of the images becomes a 1 by (M-1) matrix as shown in Equation (3.5):

$$D = [D_1, D_2, D_3, \dots, D_{M-1}] \quad (5)$$

The deviation between the base image and the deviated images was calculated by obtaining the difference between the base coordinates V_b and D_j as shown in Equation (3.6).

$$\phi = [\{V_b - D_1\}, \{V_b - D_2\}, \{V_b - D_1\}, \dots, \{V_b - D_{M-1}\}] \quad (6)$$

From the deviation (ϕ), the images were sorted in order of maximum deviation from the base image from which the most deviated images will be selected for training. The FIDE-ISA algorithm is shown in Figure 3.

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

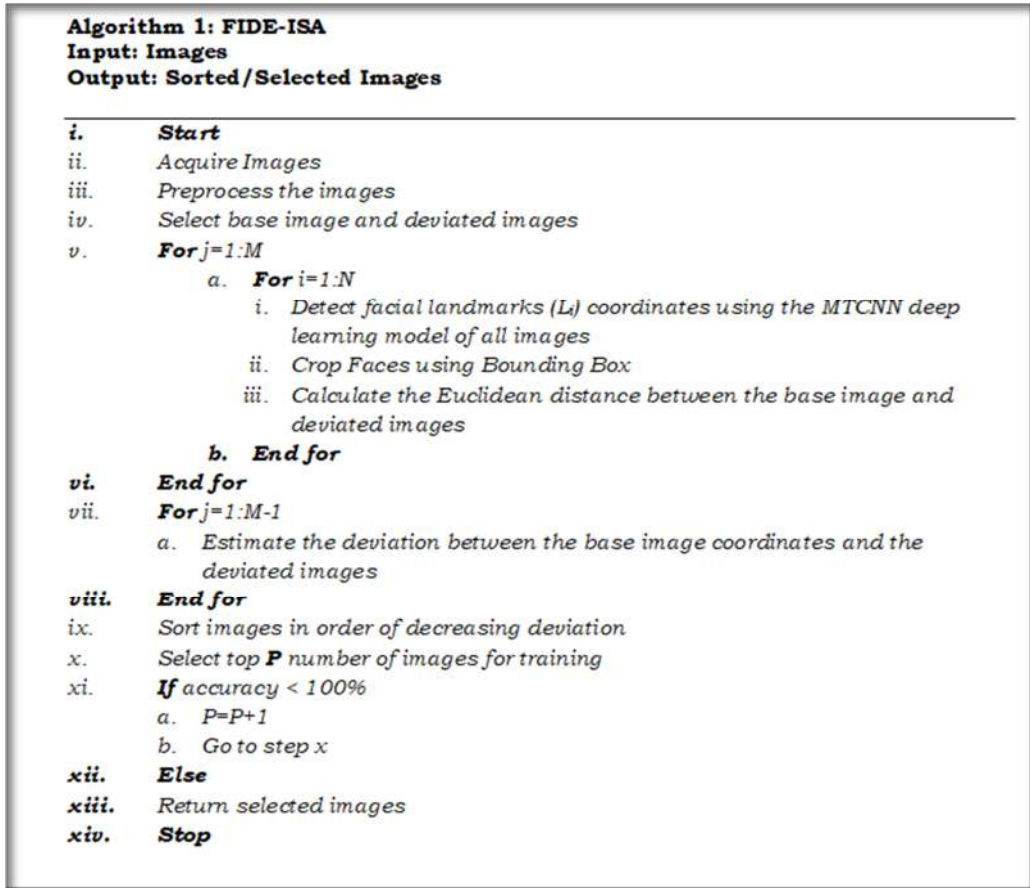


Figure 3: FIDE-ISA Algorithm

Experimentation

In this study, the hardware and software that will be used for the proposed designed CNN face recognition model are presented in Table 1. MATLAB R2020a software will be used to train and test the model, and perform performance evaluation. The experiment was carried out using an available online face dataset (faces94). The experiment was conducted by first loading the images, pre-process the image, select training samples using FIDE-ISA, extract features using AlexNet CNN for training. The trained model will then be tested by loading test images, classify the images using Softmax transfer learning and

calculate model performance as described in the flowchart in Figure 4.

Nine (9) sub-experiments were performed to determine the effectiveness of the image selection algorithm on the performance of the system. Nine different models were created using an incremental number of training samples. 18 individuals were selected comprising of 20 images per individual. The last three images from each individual were combined to form the test dataset. The first 17 images per individual make the training images. For each individual, the algorithm selects the appropriate images for training starting from 2, 3, 5, 7, 9, 11, 13, 15 and 17 respectively.

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

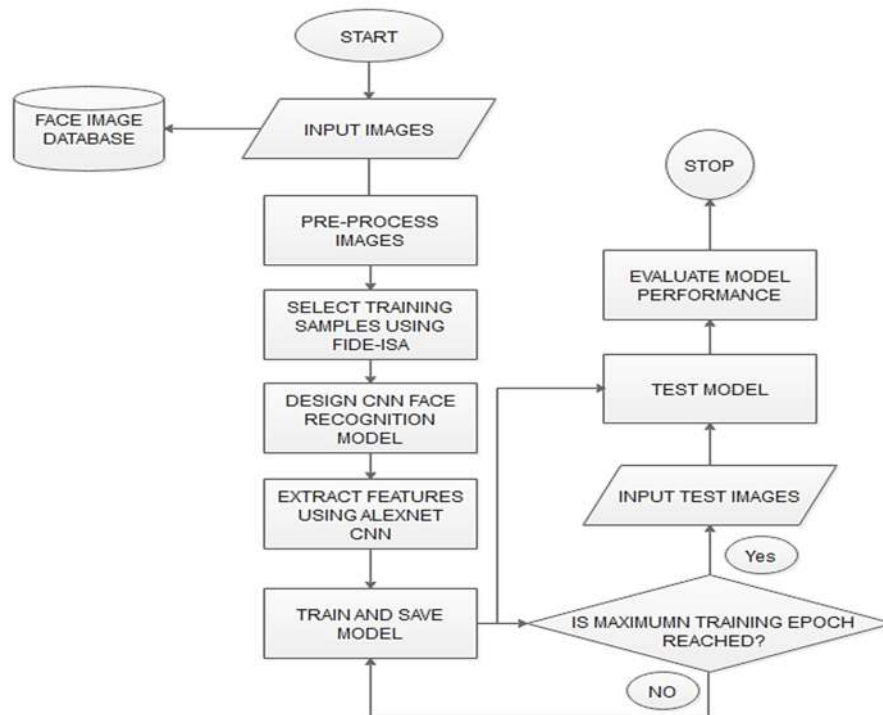


Figure 4: Flow chart of CNN Model Development Process

For Experiment 1, the system was trained with 36 faces comprising of two faces per individual. Experiment two is a model trained with 54 faces, three faces per individual, while experiment three was a model trained with 90 faces comprising of 5 faces per individual. Experiments four, five, six, seven and eight have their models trained with 126, 162, 198, 234 and 270 faces respectively comprising of 7, 9, 11, 13 and 15 faces per individual. Lastly, in experiment nine, the system was trained with 306 faces comprising 17 faces per individual. Each model developed was evaluated by measuring the systems accuracy, time of convergence and rate of convergence. The best model was selected based on the minimum training time and highest accuracy achieved. The other databases will also be trained and evaluated accordingly. The experiment was implemented using MatLab R2020a in a system with a 2.6 GHz processing speed.

Performance Evaluation

The developed face detection and recognition system were evaluated using accuracy, time of training and iteration rate. The accuracy measures the total number of correctly recognized faces against all the faces and it is given by Equation 7:

$$Accuracy = \frac{CRF}{TNF} \quad (7)$$

These metrics was adopted by most researchers in the field of face recognition (Ma, et al., 2019). The time metrics were used to measure the time complexity of the system (in seconds), which will be one of the bases for selecting or rejecting a model. The number of iteration per epoch was also measured for each model to ascertain the relationship between a number of training samples, accuracy and convergence time.

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

RESULTS AND DISCUSSION

Training Results

Table 1 shows the systems training convergence results for experiment 1 to 9. The results comprise convergence time and the convergence rate for all the experiments. From Table 1, when the number of training images is 36, the time of convergence is 108s, training with 54 images increases the convergence time to 149s, training with 90 images converges at 154s, and with 126 images, it converges at 169s. Also,

training with 162, 198, 234, 270 and 306 images converges at 331s, 455s, 484s, 600s and 664s respectively. Similarly, the number of iterations required for convergence in experiment 1, 2, 3, and 4 are 20, where the maximum images per individual are 7. For experiment 5, with 9 images per individual, the convergence rate rose to 40, while, the convergence rate increases to 60 for experiments 6 and 7 with 11 and 13 individuals per image respectively. Finally, in experiment 8 and 9, the convergence rate rose to 80.

Table 1: Training convergence result

Experiments	Number of Training Images	Images per Person	Convergence Time (s)	Convergence Rate
1	36	2	108	20
2	54	3	149	20
3	90	5	154	20
4	126	7	169	20
5	162	9	331	40
6	198	11	455	60
7	234	13	484	60
8	270	15	600	80
9	306	17	664	80

The results obtained during the training phase of the developed face recognition system for experiments 1 to 9 are analyzed in two ways. The first is the analysis of the number of training images and time of convergence analysis shown in Figure 5 and the second is the analysis of the number of image per individual, the convergence

time and convergence rate shown in Figure 6. From Figure 5, as the number of training samples increases (the pink curve), the time of convergence (the green curve) also increases. The increase in convergence time from the fifth experiment to the ninth experiment exceeds the increase from the first to the fourth iteration.

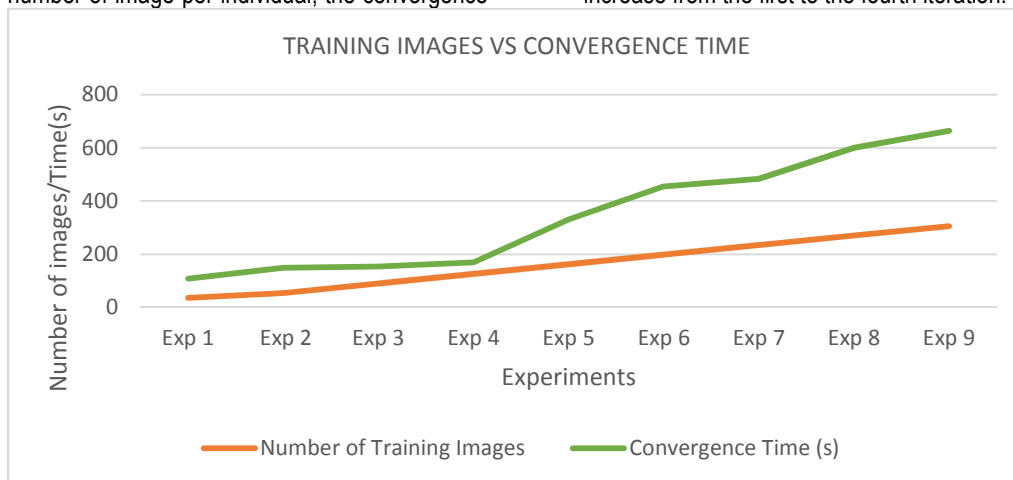


Figure 5: Training Convergence Time Analysis

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved

This is attributed to the fact that as the number of images exceeds 300, the number of epoch per iteration increases to 2. Similarly, as the number of images reaches 600, the epoch per iteration increases to 4. This lead to a consequential increase in the time taken to train the system. Figure 6 shows three curves, the yellow curve represents the number of images per individual, the pink curve represents the time of convergence in minutes and the blue curve represents the convergence rate. From experiment 1 to experiment 4, the number of images per individual is 2, 3, 5, and 7 respectively,

resulting in a single epoch per iteration with a convergence iteration of 20. The fifth experiment was trained with 9 images per individual which converges in 40 iterations with 2 epoch per iteration. Experiments 6 and 7 have 11 and 13 images per individual and converges in 60 iterations with 3 epoch per iteration. Experiments 8 and 9 converge in 80 iterations with 4 epoch per iteration. These results indicated that the number of epoch per iteration is proportional to the number of images per individual used to train the system which also affects the training time.

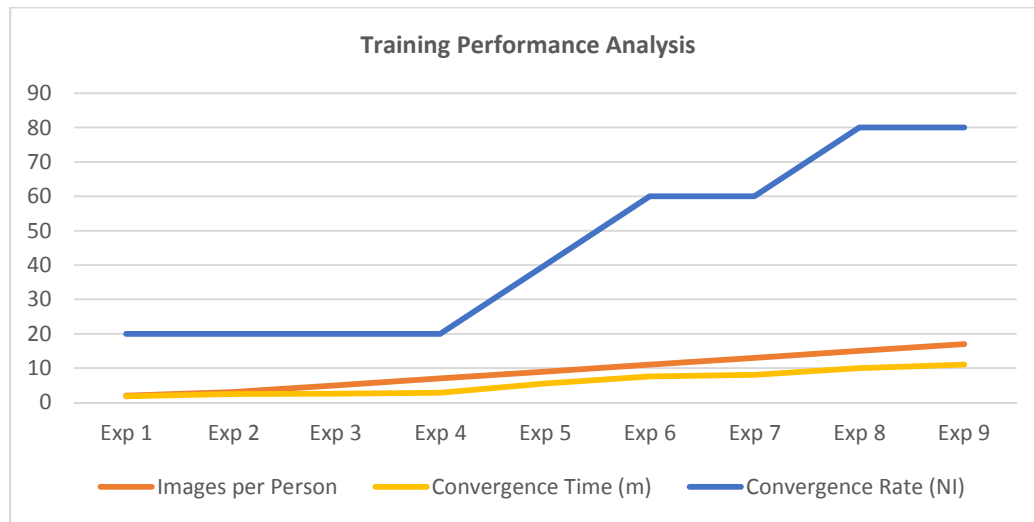


Figure 6: Training Performance Analysis

Testing Results

Three images from each individual not used for training were used for testing the system for each experiment. The 54 test images were checked to determine if they belong to a particular individual or not. Table 2 shows the summary of the numbers of correctly classified and wrongly

classified faces for experiments 1 to experiments 9. From the table, 51 faces were correctly recognized and 3 were wrongly recognized for experiment one. For experiment 2 to experiment 9, 54 faces were correctly and recognized and none was wrongly recognized.

Table 2: Summary of test recognition results

Experiments	Number of Correctly Recognized Faces	Number of Wrongly Recognized Faces
1	51	3
2	54	0
3	54	0
4	54	0
5	54	0

*Corresponding author: Hassan, O.S. alhassan77077@gmail.com Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved



Experiments	Number of Correctly Recognized Faces	Number of Wrongly Recognized Faces
6	54	0
7	54	0
8	54	0
9	54	0

Test Performance Evaluation and Analysis

Table 3 shows the performance evaluation results of the system testing for the 9 experiments. The accuracy of the system for experiment 1 is 94.44%, while the accuracy for experiments 2 to 9 is 100% respectively. The result shows an increase in the number of training

image (red bar) 36 to 54 images improves overall performance from 94.44% to 100% (green bar). Further increase of the training samples from 54 to 90, 90 to 126, 126 to 162, 162 to 198, 198 to 234, 234 to 270, and 270 to 306 (the red bar) all produces 100% accuracy as shown in the green bar in Figure 7.

Table 3: Test Performance evaluation

Experiments	Number of Training Images	Convergence Time (s)	Accuracy (%)
1	36	108	94.44
2	54	149	100
3	90	154	100
4	126	169	100
5	162	331	100
6	198	455	100
7	234	484	100
8	270	600	100
9	306	664	100

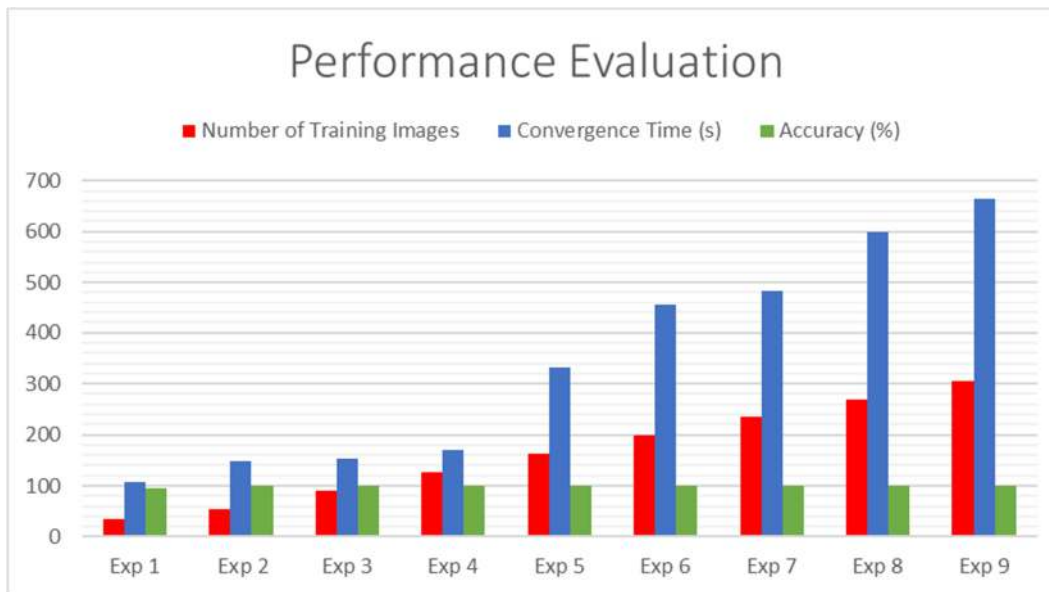


Figure 7: Test Performance Analysis

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved



From Figure 7, it can be deduced that experiment two with 54 training images, 3 images per individual produces the most desirable results with 100% accuracy achieved in 149 seconds. This result indicates that training a face recognition model with a fewer number appropriate images produces the same result in terms of accuracy in a much faster time. The proposed algorithm and face recognition model proposed in this research work successfully reduced the training time from 664s to 149s. This results in 345.64% training time reduction which is very significant in reducing the time spent in training face recognition systems.

CONCLUSION

The training result of the developed system for the nine experiments performed indicated that the time of convergence is proportional to the number of training samples or a number of images per individual selected for training. The convergence rate, which is the number of epoch per iteration is also proportional to the number of images per individual. Training with less than eight images per individuals required only one epoch per iteration to converge. Training with 9 images per individual requires two epochs per iteration, while 11 to 13 images per individual results in three epochs per iteration. Finally, above 13 images per individual, required four epochs per iteration. For all the models trained with the increasing number of training samples, the convergence accuracy is 100%. This indicated that selecting the appropriate type of images based on the deviation of the facial features produces the same result in terms of accuracy when compared with models trained with a higher number of images.

In terms of testing the developed trained system, the results obtained indicated that models trained with 54 images and above produce 100% accurate authentication of faces. Only the model trained with less than 54 images produces less than 100% accurate authentication. This is an indication that training a facial recognition system with few images can produce better results when it comes to accuracy and time of convergence. This research has contributed immensely to knowledge by developing an algorithm that selects appropriate and few images for training a facial

recognition system thereby enhancing the training and testing phases of a facial recognition system.

REFERENCES

- Almabdy, S., & Elrefaei, L. (2019). Deep convolutional neural network-based approaches for face recognition. *Applied Sciences (Switzerland)*, 9(20). <https://doi.org/10.3390/app9204397>
- Alqadi, Z., Khrisat, M., Dwairi, M., & Khawatreh, S. (2020). *Building Face Recognition System (FRS)*. 9(June), 15–24.
- Arafah, M., Achmad, A., Indrabayu, & Areni, I. S. (2020). Face Identification System Using Convolutional Neural Network for Low-Resolution Image. *2020 IEEE International Conference on Communication, Networks and Satellite, Comnetsat 2020 - Proceedings*, 55–60. <https://doi.org/10.1109/Comnetsat50391.2020.9328967>
- Coskun, M., Ucar, A., Yildirim, O., & Demir, Y. (2017). Face recognition based on convolutional neural network. *Proceedings of the International Conference on Modern Electrical and Energy Systems, MEES 2017, 2018-Janua*(January 2018), 376–379. <https://doi.org/10.1109/MEES.2017.8248937>
- Filippidou, F. P., & Papakostas, G. A. (2020). Single Sample Face Recognition Using Convolutional Neural Networks for Automated Attendance Systems. *4th International Conference on Intelligent Computing in Data Sciences, ICDS 2020*. <https://doi.org/10.1109/ICDS50568.2020.9268759>
- Ling, H., Wu, J., Huang, J., Chen, J., & Li, P. (2020). Attention-based convolutional neural network for deep face recognition. *Multimedia Tools and Applications*, 79(9–10), 5595–5616. <https://doi.org/10.1007/s11042-019-08422-2>
- Marsot, M., Mei, J., Shan, X., Ye, L., Feng, P., Yan, X., Li, C., & Zhao, Y. (2020). An adaptive pig face recognition approach using Convolutional Neural Networks. *Computers and Electronics in Agriculture*,

*Corresponding author: Hassan, O.S. ✉ alhassan77077@gmail.com ✉ Department of Computer Science, University of Abuja. © 2021 Faculty of Technology Education, ATBU Bauchi. All rights reserved



173.
<https://doi.org/10.1016/j.compag.2020.105386>
- Nagpal, S., Singh, M., Singh, R., & Vatsa, M. (2019). Deep learning for face recognition: Pride or prejudiced? *ArXiv, July*.
- Ogbuju, E., Adetayo, A. P., & Obilikwu, P. (2020). A Face Recognition System for Attendance Record in a Nigerian University. *Journal of Scientific Research and Development, 19*(June), 38–45.
- Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). *Deep Face Recognition. Section 3*, 41.1-41.12. <https://doi.org/10.5244/c.29.41>
- Pei, Z., Xu, H., Zhang, Y., Guo, M., & Yee-Hong, Y. (2019). Face recognition via deep learning using data augmentation based on orthogonal experiments. *Electronics (Switzerland), 8*(10), 1–16.
<https://doi.org/10.3390/electronics8101088>
- Peng, X., Ratha, N., & Pankanti, S. (2016). Learning face recognition from limited training data using deep neural networks. *Proceedings - International Conference on Pattern Recognition, 0*, 1442–1447.
<https://doi.org/10.1109/ICPR.2016.7899840>
- Swapna, M., Sharma, Y. K., & Prasad, B. M. G. (2020). A Survey on Face Recognition Using Convolutional Neural Network. *Advances in Intelligent Systems and Computing, 1079*, 649–661.
https://doi.org/10.1007/978-981-15-1097-7_54
- Tao, K., He, Y., & Chen, C. (2019). Design of Face Recognition System Based on Convolutional Neural Network. *Proceedings - 2019 Chinese Automation Congress, CAC 2019*, 5403–5406.
<https://doi.org/10.1109/CAC48633.2019.8996236>