A fast and non-trainable facial recognition system for schools

Kazeem Oyebode[1], Kingsley Chiwuike Ukaoha[2]
[1]Design and New Media Department, Pan-Atlantic University, Lagos, Nigeria
[2]Department of Computer Science, University of Benin, Benin, Nigeria

Article Info

Article history:

Received Sep 1, 2021
Revised Nov 6, 2021
Accepted Dec 2, 2021

Keywords:

Deep learning
Facial recognition
Feature vector
Non-trainable system
Pattern matching

ABSTRACT

Deep learning models have been at the forefront of facial recognition because they deliver improved classification accuracy over traditional ones. Regardless, deep learning models require an extensive dataset for training. To significantly cut down on its training time and dataset volume, pre-trained models, have been used although, they are still required to undergo the usual training process for custom facial recognition tasks. This research focuses on an improved facial recognition system that lacks the training and retraining requirements. The system uses an existing deep learning feature extraction model. First, a user stands before a camera-enabled system. After that, the user supplies a unique identification number to fetch a corresponding face image from the database. This process generates two face feature vectors. One from the camera and that retrieved from the database. The cosine distance function determines the similarity value of these vectors. When the cosine distance value falls below a set threshold, the face is recognized and access granted. If the cosine distance of the two vectors gives a value above this threshold, access is denied. The proposed model performs satisfactorily on publicly available datasets.

Corresponding Author:

Kazeem Oyebode
Design and New Media Department, Pan-Atlantic University
Km 52 Lekki - Epe Expy, 105101, Lagos, Nigeria
Email: kazeemkz@gmail.com

1. INTRODUCTION

Facial recognition systems have been beneficial in resolving a host of problems such as identity verification [1], [2], facial expression identification [3], facial emotions classification [4], crime resolution [5], vehicle security [6], [7] and host of others. The rate of kidnappings in Nigerian schools has increased owing to a lack of security frameworks. Visitors can gain access without proper checks. Insecurity in schools could have devastating effects on the girl child education [8], [9] and could impact their academic performances [10]. According to a source in [11], about 13.5 million children are out of school owing to insecurity. Facial recognition systems could help mitigate this challenge.

One of such is the local binary pattern (LBP) [12], [13]. The LBP uses the binary representation to represent the texture properties of face images. An image is sub-divided into cells (say 20 by 20 pixels) then each pixel's neighboring properties are examined. A comparison between each pixel *r* in a considered cell and its eight surrounding pixels is carried out. Where *r* intensity value is greater than a surrounding pixel *v*, *v* then holds a binary value zero otherwise, *v* holds value one. The binary values of the 8-pixel neighbors of *r* are collected and transformed into a sequence of numbers. The number combinations from individual cells over the entire image are projected into a histogram of features. Also, a transformation is carried out on the face to be recognized (g) into a histogram of features. A comparison is made between the feature histogram of the training dataset and that of g. The Euclidean distances of *g* to various histograms of the training dataset are juxtaposed.

The smallest Euclidean distance to g is considered the recognized face [12]. The principal component analysis (PCA) [14], [15] is another facial recognition model and works by maximizing the variance amongst classes of data. PCA extracts unique features from a collection of images and projects them onto a given face space. For frontal face recognition, the PCA has the advantage of its rapid computation [14]. The fishers linear discriminant analysis (LDA) [16], [17] is another model that shares a similar prospect with the PCA for facial recognition. It maximizes the distances amongst data clusters while ensuring that each group is tightly packed.

Recently, the availability of cheap yet powerful computing capability has made deep learning models attractive for solving facial recognition problems. In addition, its accuracy level significantly outperforms the models previously discussed. However, they have similar characteristics- they both undergo training. More importantly, the convolution neural network (CNN) [18], [19] needs more training datasets for its model to perform optimally [20]. Its heavy reliance on lots of training images has also been solved partly by image augmentation with the aim of improved performance [21]. This development could be perceived as a downside although, pre-trained models mitigate the reliance on large training datasets. For example, in [22], the pre-trained model is used with a face classification layer. This approach provides the possibility of using only a few training images in a deployed environment. In a scenario where the faces to be recognized are known in advance, this approach would work well. However, in an environment that attracts all kinds of new faces that need to be granted access into a facility, this approach in [22] may not be appropriate. The reason is that the model has to undergo a retraining process on new faces. This development makes its deployment impractical in real-life scenarios such as school environments where visitors require screening in real-time to gain access to school facilities.

2. THE PROPOSED METHOD

To mitigate this challenge of training and retraining facial recognition systems, we propose a model anchored on a CNN architecture seen in [23]. This architecture is trained on 2.6 million face images. One advantage of this architecture is that it can be adapted to generate feature vectors of faces at its output layer when a face image is passed into its input layer. The proposed workflow has three stages. The first is the facial registration stage. Here, only a single photo of a student or a visitor is captured and stored with a student or visitor's identity number in a central database. Secondly, in the recognition stage, the human before the camera supplies the identification number. This identification number serves to retrieve a matched face stored in the database. Thirdly, a feature vector I is generated from the faces retrieved via the database. In addition, another feature vector named K is generated from the photo captured via the access control camera. A comparison is made between these feature vectors using a cosine similarity measure [24]. A threshold value of 0.4 classifies the input photo provided by the camera. At a distance, not more than 0.4, the model returns 'the face presented is recognized'. In addition, if the model returns a value of more than 0.4, an output 'the face presented is not recognized' is generated. Figure 1 exemplifies the registration process.

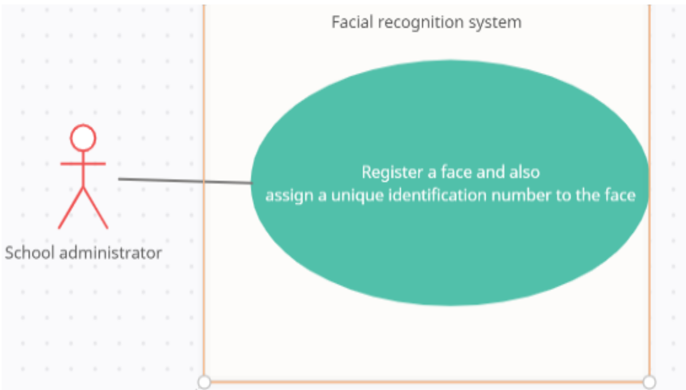

Figure 1. Face capturing phase

Figure 2 details the facial recognition use case. This process requires that the user enters an identification number into the system. After that, the identification serves to retrieve a matched face in the database. The feature vector of the face retrieved is then compared with that fetched from the camera. If there is a match, the system grants access if otherwise, access is allowed. This process is fast and practical as it requires no training and a retraining process. Algorithm 1 also shows the inner workings of the proffered model.

Figure 3 provides further insight into various components of the proffered algorithm. The system functionalities are in two parts-they are the front end and the backend. The front end has two units- they are the admin unit and the user unit. The admin unit enables students/visitors registration while the user unit facilitates the authentication process. The backend saves user data and also converts faces into feature vectors. The cosine distance function then compares these vectors for a match.

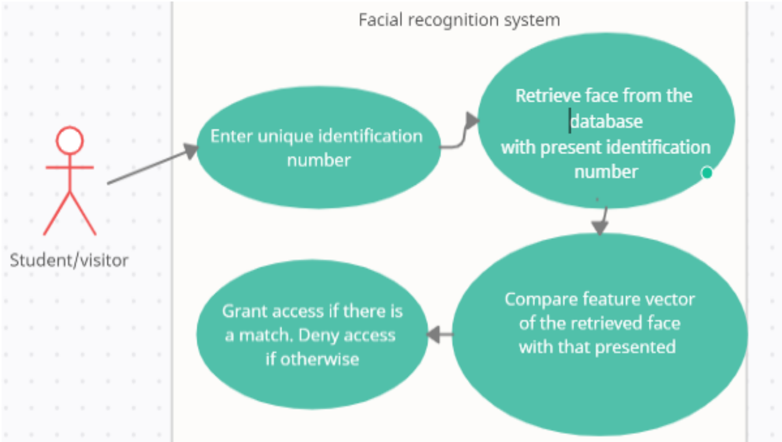

Figure 2. Facial recognition phase

Algorithm 1. A fast facial recognition access control system for schools

```
Inputs: F, it represents the presented face before a camera
Outputs: O represents system output 'face recognized' or 'face not recognized'
Precondition: A replica of Face F must be registered in the database with an
identification number N
Begin:
Initialize
Retrieve face from the camera and assign to F.
Convert F into a single feature vector I
The user also enters the identification number and assigned to N
Retrieve face V using N from the database.
Convert V into a single feature vector K
G = cosine_distance (I,K)
If G >0.4
    O = "face not recognized"
Else
    O = 'face recognized'
return O
```

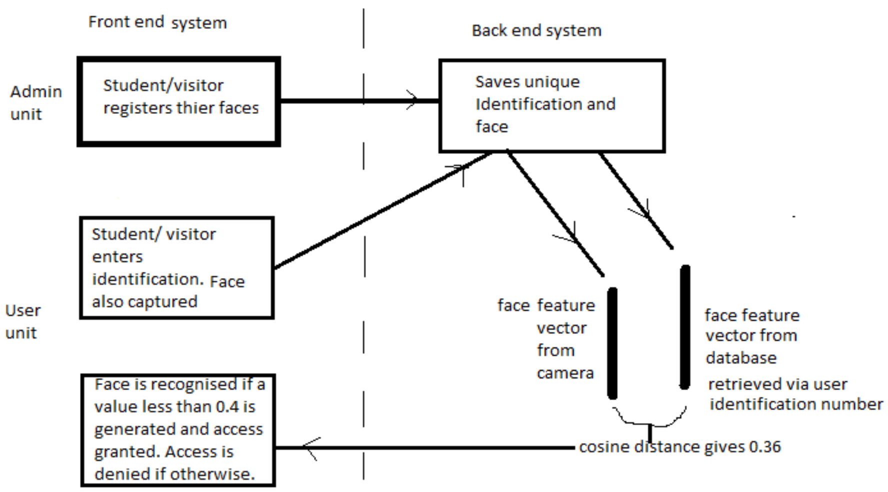

Figure 3. The architecture of the proposed algorithm

3. RESEARCH METHOD

The proposed model is tested on two publicly available frontal face datasets using an i7 laptop processor with 16GB of random access memory (RAM). Two publicly available datasets (Face95 and face96) [25], provided by the University of Essex are used for the evaluation phase. Face95 has 72 unique faces (72 persons), and each face has 20 frontal images. There Face95 has a total of 1,440 faces. Of these images, only 72 images were held back (one picture per person) as images stored in the database. For face96 dataset, it has 151 unique faces (151 persons). Each unique face has 20 samples. Therefore, Face96 has a total of 3020 faces. Of these images, only 151 images were uploaded into the database. The accuracy metric in 1 evaluates the proposed model.

$$A\ (\%) = \frac{Total\ no.of\ correctly\ classified\ images}{Totla\ no.of\ images\ referenced} \quad (1)$$

4. RESULTS AND DISCUSSION

Tables 1 and 2 show the performances of the proposed model. We observed that it performed better than all considered models. Although it only had a reference image per person in terms of face feature vector comparison, it performed better than others. This performance is expected since the development of the offered model sits on a robust feature extraction model. Other models are sensitive to noise, for example, PCA and LDA. The local binary pattern (LBP) has less light sensitivity, consequently performs better.

To further demonstrate the robustness of the proposed architecture, we compare a picture taken in 2017 (as the database face) with that extracted from the webcam in 2021. Figure 4 depicts the photo taken in 2017 used as the reference image stored in the database by the administrator, and an identification number of 011075 is attached to it. This number could also serve as a student's matriculation number. This process is a one-time registration process. The next is the recognition stage seen in Figure 5, where the student would enter the matriculation number. After that, the face feature vector, extracted from the camera as shown in Figure 6, is compared to that stored in the database, using a cosine function. Finally, if the cosine distance function gives a value less than 0.4 as shown in Figure 7, the face is indeed recognized. This development also shows that an old picture of a student stored in the first year in school can be relevant in granting access for several years. In addition, it shows that distinct facial features do not fade over time. The facial recognition phase lasts for only five seconds. Figure 8 shows what might happen if another person's identification is entered as seen in Figure 9. The system will deny access as a cosine distance returns a value more than 0.4.

Table 1. Models' performance on face95

Model	A (%)
PCA	80.3
LDA	92.7
LBP	89.8
Proposed	99.7

Table 2. Models' performances on face96

Model	A (%)
PCA	62.1
LDA	77.5
LBP	82
Proposed	98.6

Figure 4. Reference image stored in the database by the administrator (taken in 2017). An identity number of 011075 is attached to this image

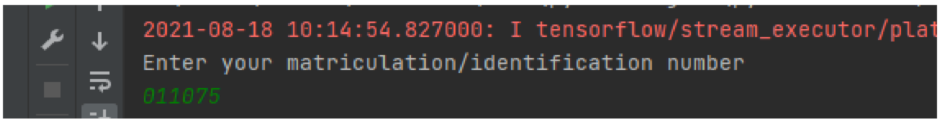

Figure 5. The user enters the identification number in the system

Figure 6. Image captured by the system via the webcam

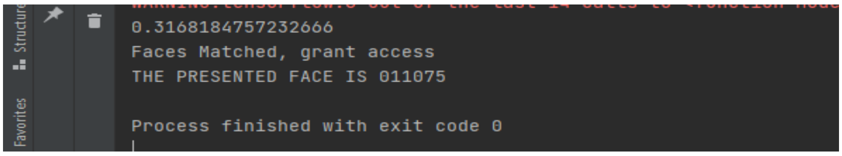

Figure 7. The cosine distance function compares feature vectors and returns a value of 0.3168. This value is less than the set threshold value of 0.4

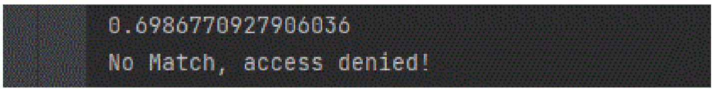

Figure 8. The cosine distance function compares feature vectors and returns a value of 0.6986. This is more than the set threshold value of 0.4

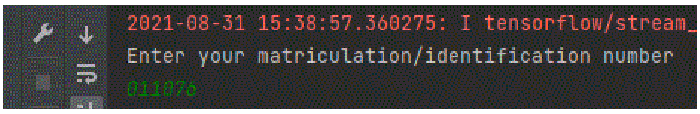

Figure 9. The user enters a wrong identity number of 011076

5. CONCLUSION

Facial recognition systems are crucial for identification purposes and can play a critical role in granting access to facilities, especially in schools. Existing models such as deep learning often require training and retraining. This approach may work in environments where the faces to be recognized are known in advance. Where faces to be recognized are not known in advance, the proposed model can be relevant, especially in schools. The model leverages identities assigned to persons for the retrieval of stored faces. This development makes it possible to compare the feature vector of the retrieved face to that extracted from the webcam or camera. The proffered solution is tested on two face datasets with promising results. We hope that it will be installed for widespread use in Nigeria and beyond. However, financing is still a big challenge as the initial set-up costs of the components could be high.

REFERENCES

[1] B. R. Ilyas, B. Mohammed, M. Khaled, and K. Miloud, "Enhanced Face Recognition System Based on Deep CNN," in *Enhanced Face Recognition System Based on Deep CNN6th International Conference on Image and Signal Processing and their Applications (ISPA)*, 24-25 Nov. 2019, pp. 1-6, doi: 10.1109/ISPA48434.2019.8966797.

[2] L. Yuan, Z. Qu, Y. Zhao, H. Zhang, and Q. Nian, "A convolutional neural network based on TensorFlow for face recognition," 2017," in *IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chongqing, China. 25-26 March 2017, pp. 525-529, doi: 10.1109/IAEAC.2017.8054070.

[3] Y. Wang, Y. Li, Y. Song, and X. Rong, "The Application of a Hybrid Transfer Algorithm Based on a Convolutional Neural Network Model and an Improved Convolution Restricted Boltzmann Machine Model in Facial Expression Recognition," *IEEE Access*, vol. 7, pp. 184599-184610, 2019, doi: 10.1109/ACCESS.2019.2961161.

[4] A. Atanassov and D. Pilev, "Pre-trained Deep Learning Models for Facial Emotions Recognition," in *International Conference on Automatics and Informatics (ICAI)*, Varna Bulgaria, 2020, pp. 1-6, doi: 10.1109/ICAI50593.2020.9311334.

[5] B, Orellana, L, Álvarez, and J. Armas-Aguirre, "Face Recognition for Criminal Identification," in *Proceedings of the 6th Brazilian Technology Symposium (BTSym'20). BTSym 2020. Smart Innovation, Systems and Technologies*, 2020.

[6] A. A. Elngar and M. Kayed, "Open Computer Science," *Vehicle Security Systems using Face Recognition based on Internet of Things*, vol. 10, no. 1, pp. 17-29, 2020, doi: 10.1515/comp-2020-0003.

[7] R. F. Rahmat, M. P. Loi, S. Faza, D. Arisandi, and R. Budiarto, "Facial Recognition for Car Security System Using Fisherface Method," in *Journal of Physics: Conference Series*, vol. 1235, no. 1, 2019, doi: 10.1088/1742-6596/1235/1/012119.

[8] O. Sanni, "Effects of insecurity and challenges on females' education in Nigeria," *African Journal for the Psychological Study of Social Issues,* vol. 18, no. 3, 2015.

[9] U. Bassey, "Insecurity and girl-child education in Nigeria," *European Journal of Education Studies,* vol. 2, no. 11, 2016, doi: 10.46827/ejes.v0i0.405.

[10] M. Ojukwu, "Effect of Insecurity of School Environment on the Academic Performance of Secondary School Students in Imo State," *International Journal of Education & Literacy Studies,* vol. 1, no. 5, 2017, doi: 10.7575/aiac.ijels.v.5n.1p.20.

[11] T. Cable, "The Cable," [Online]. Available: https://www.thecable.ng/insecurity-is-crippling-nigerias-education-system-heres-what-we-must-do (accessed Aug. 27, 2021).

[12] S. Zhu and J. Feng, "A Novel Facial Appearance Descriptor Based on Local Binary Pattern," in *Chinese Conference on Pattern Recognition*, Beijing, China, 22-24 Oct. 2008, pp. 1-4, doi: 10.1109/CCPR.2008.58.

[13] N. N. Fatihah, G. Ariyanto, A. J. Latipah, and D. M. Pangestuty, "Face Recognition Using Local Binary Pattern and Nearest Neighbour Classification," in *International Symposium on Advanced Intelligent Informatics (SAIN)*, Indonesia, 29-30 Aug. 2018, pp. 142-147, doi: 10.1109/SAIN.2018.8673375.

[14] R. Kaur and E. Himanshi, "Face recognition using Principal Component Analysis," in *IEEE International Advance Computing Conference (IACC)*, Banglore, 2015, pp. 585-589, doi: 10.1109/IADCC.2015.7154774.

[15] W. Deng, J. Hu, J. Lu, and J. Guo, "Transform-Invariant PCA: A Unified Approach to Fully Automatic FaceAlignment, Representation, and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 36, no. 6, pp. 1275-1284, 2014, doi: 10.1109/TPAMI.2013.194.

[16] M. Bober, K. Kucharski, and W Skarbek, "Face Recognition by Fisher and Scatter Linear Discriminant Analysis," *Computer Analysis of Images and Patterns, 10th International Conference*, Groningen, The Netherlands, August 25-27, 2003, doi: 10.1007/978-3-540-45179-2_78.

[17] W. Yunzhu and C. Yunli, "A new feature extraction algorithm based on fisher linear discriminant analysis," in *the 3rd International Conference on Control, Automation and Robotics (ICCAR)*, 24-26 April, Nagoya, Japan 2017, pp. 414-417, doi: 10.1109/ICCAR.2017.7942729.

[18] S. Albawi, T. A. Mohammed, and S. Al-aZawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, Antalya, Turkey, 21-23 Aug. 2017 pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[19] T. Zebin, P. J. Scully, N. Peek, A. J. Casson, and K. B. Ozanyan, "Design and Implementation of a Convolutional Neural Network on an Edge Computing Smartphone for Human Activity Recognition," *IEEE Access*, vol. 7, pp. 33509-133520, 2019.

[20] D. Wang, H. Yu, D. Wang, and G. Li, "Face Recognition System Based on CNN," in *International Conference on Computer Information and Big Data Applications (CIBDA)*, Guiyang, China, 17-19 April 2020, pp. 470-473, doi: 10.1109/CIBDA50819.2020.00111.

[21] L. Peng, S. Baoye, and X Lin, "Human face recognition based on convolutional neural network and augmented dataset," *Systems Science & Control Engineering,* vol. 9, no. 2, pp. 29-37, 2021, doi: 10.1080/21642583.2020.1836526.

[22] R. M. Prakash, N. Thenmoezhi, and M. Gayathri, "Face Recognition with Convolutional Neural Network and Transfer Learning International Conference on Smart Systems and Inventive Technology (ICSSIT)," in *International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, 27-29 Nov. 2019, pp. 861-864, doi: 10.1109/ICSSIT46314.2019.8987899.

[23] M. P. Omkar, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *British Machine Vision Conference*, 2015. doi: 10.5244/C.29.41.

[24] S. Ontañón, "An overview of distance and similarity functions for structured data," *Artif Intell,* vol. 53, pp. 5309-5351, 2020, doi: doi.org/10.1007/s10462-020-09821-w.

[25] S. Libor, "Face Recognition Database," [Online]. Available: https://csperson.kku.ac.th/chakchai/faceDBweb/face_dataset.html (accessed Jan. 2, 2018).

BIOGRAPHIES OF AUTHORS

Kazeem Oyebode received the Ph.D degree in image processing and computer vision from the University of KwaZulu-Natal, South Africa, in 2017. He currently lectures at Pan-Atlantic University, Lagos. Nigeria. He can be contacted at email: kazeemkz@gmail.com.

Kingsley Chiwuike Ukaoha is an Associate Professor of Computer Science and the Deputy Director (ICT), Centre for Distance Learning at the University of Benin, Nigeria. He can be contacted at email: kingsley.ukaoha@uniben.edu.