



# INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : [www.joiv.org/index.php/joiv](http://www.joiv.org/index.php/joiv)



## Real-time Triplet Loss Embedding Face Recognition for Authentication Student Attendance Records System Framework

Hady Pranoto<sup>a,\*</sup>, Oktaria Kusumawardani<sup>a</sup>

<sup>a</sup> Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia

Corresponding author: \*hadypranoto@binus.ac.id

**Abstract**— The number of times students attend lectures has been identified as one of many success factors in the learning process in many studies. We proposed a framework of the student attendance system by using face recognition as authentication. Triplet loss embedding in FaceNet is suitable for face recognition systems because the architecture has high accuracy, quite lightweight, and easy to implement in the real-time face recognition system. In our research, triplet loss embedding shows good performance in terms of the ability to recognize faces. It can also be used for real-time face recognition for the authentication process in the attendance recording system that uses RFID. In our study, the performance for face recognition using k-NN and SVM classification methods achieved results of 96.2 +/- 0.1% and 95.2 +/- 0.1% accordingly. Attendance recording systems using face recognition as an authentication process will increase student attendance in lectures. The system should be difficult to be faked; the system will validate the user or student using RFID cards using facial biometric marks. Finally, students will always be present in lectures, which in turn will improve the quality of the existing education process. The outcome can be changed in the future by using a high-resolution camera. A face recognition system with facial expression recognition can be added to improve the authentication process. For better results, users are required to perform an expression instructed by face recognition using a database and the YOLO process.

**Keywords**— Computer vision; face recognition; recording attendance; framework.

Manuscript received 25 Nov. 2020; revised 2 Jan. 2021; accepted 5 Mar. 2021. Date of publication 30 Jun. 2021.  
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



### I. INTRODUCTION

In several studies, one of many success factors in the academic process is the amount of how many times students attend the lectures. The present level of students in lecture sessions can increase the level of absorption and dissemination of the course given by lecturers. The level of absorption and dissemination of the course influences the success rate of student studies, proved by research conducted by Fadelelmoula [1], a student with a high rate of attendance had better academic performance and a high score test. Tetteh [2] found that the attendance rate and learning time amount significantly affect the study's outcome. Bartanen [3] shows that the attendance rate of students in a course increases academic performance besides how the teacher delivers the course. After conducting several observations and studies, we conclude that the attendance rate may affect the performance and success in academic fields. However, the problem is how to make the students have consistency in attending the lecture sessions by providing a robust attendance system which is expected to be difficult to be faked and falsified.

In many universities, student attendance rates at lectures session are very low. It results in low student academic performance. To increase student attendance rate, we need a smart attendance system, a difficult system to manipulate. This system can ensure the presence of students in a class, in this case: the authentication process. An authentication process can be done in various ways, such as sign-on paper or by using Radio Frequency Identification (RFID) card as an identity. RFID can easily be manipulated and falsified by bringing the other friend's card into class. For that reason, we need another method to make the authentication process in the attendance system difficult to be manipulated or falsified. One method of authentication which is difficult to falsify is by utilizing biometric marks from each student. Biometric identification, such as using fingerprints, retina, and face, is quite difficult to be manipulated. They are very individual-based. Each method of authentication has advantages and disadvantages. Authentication using fingerprints or retina has a higher level of security against falsified, yet a high cost is required to implement this method. Another option is to use face recognition which using biometric marks that are on the

face. Zang et al. [4] research demonstrated that facial recognition is good for the authentication method. Thus, the author decides to implement facial recognition in the student attendance recording system to change students' behavior, so that they can always attend the lessons. It can also provide an increased level of academic teaching and learning systems into an intelligent level system.

## II. MATERIAL AND METHOD

### A. Related Work

Face recognition as an authentication method has become popular because it has advantages over identity matching using other biometric features. Face recognition is natural and does not intrusive because face image capturing can be done in the distance. Face recognition defines as ".... a biometric method of identifying a person based on a photograph of their face; biometric methods use biological traits to identify people" [5]. In general, facial recognition using a computer is a process of pattern recognition from a face image captured by a camera. Face images can represent either 2 dimensional (2D) or 3 dimensional (3D) objects in varied lighting conditions, poses, and expressions [6].

The implementation of face recognition in daily life is utilized in several fields such as:

- Identification: driver's license card, immigration, resident card, passport, voter registration.
- Access control: border crossing control, facility access, vehicle access, kiosk, ATM, computer access, program access, computer network access, online program access, online transaction access, distance learning access, online exam access, exam access (Ye & Hu, 2017), database access, and so on [7], [8].
- Security: terrorist warnings, secure boarding system, stadium visitors scanning, computer security, computer application security, database security, file encryption, intranet or internet security, medical history (medical records), terminal stock trading security, payment method, surveillance at nuclear power plants, park monitoring.
- Law enforcement: detention of suspects, recognition of shoplifters, recognition of criminals, tracking of suspects and investigations, background checks of suspects, environmental monitoring, identification of fraud at gambling establishments, post-event analysis, social security crimes (insurance fraud/social security)[9].
- Information retrieval: indexing database by face, and searching record (retrieval), auto labelling on the face, face classification.
- Multimedia Management: face-based search, face segmentation video, event detection.
- Human-Computer Interaction: interactive game, proactive computing.
- Authentication: Smart Home, Smart Office, Smart City.
- Social security: identification of recipients of the social fund, missing person search.
- Anti-Criminal [10].

The face recognition approach has many methods so far. Some of them are using linear projection appearance-based, non-linear projection, and neural network. In a linear

projection appearance-based, features are projected into a linear subspace. Example of this method is Principle Component Analysis (PCA) [11], [9], Independent Component Analysis (ICA)[12], Linear Discriminant Analysis (LDA) [13], [14], Two Dimensional PCA (2DPCA) [15], and Linear Regression Classification (LRC) [16], Fisher Linear Discriminant [17].

Non-linear appearance-based projection approaches map face input images into high-dimensional space, and the manifold of the face is linear and simplified. Examples that using this method are kernel-PCA(KPCA), kernel-LDA [18].

Other approaches use the Local Appearance Feature: Local Binary Pattern (LBP) [19], [20], [10]. Gabor Phase Pattern[21], and Local Gabor Binary Pattern [22], Discriminative Common Vectors (DCV)[23], Support Vector Machine (SVM)[24], Neighborhood Preserving Projections (NPP)/ Orthogonal Neighborhood Preserving Projections (ONPP)[25], Combine between Kernel Discriminant Analysis(KDA), K-Nearest Neighbor and Support Vector Machine[26], a modified LBPH algorithm based on pixel neighborhood gray median(MLBPH) [27]. This method makes use of local patterns from the face and its surroundings to get unique facial features. Both Non-linear Appearance-based projection and the Local Appearance Feature approach are more accurate than the classic methods, but they are still not enough to implement in the attendance system.

The newest approach in face recognition is to use neural networks, which provide more accurate results. There are two types of neural networks: shallow neural network and deep neural network. A shallow neural network has a shallow layer in architecture, and a deep neural network has many layers in architecture to extract features. An example that uses a shallow network approach can be found in [28], and one that uses a Deep neural network can be found in Boltzmann machines [29], DeepFace [30], DeepID2 [31]. A deep neural network has high accuracy, but the approach is not easy to implement and requires a high level of computation. An approach using neural networks yet not too deep in architecture has a quite high accuracy. Xu et al. [32] proposed a 3D-aided 2D face recognition system (3D2D-PIFR) that robust to pose variation as large as 90°. Convolutional Neural Network faces recognition was also proposed by Qiao & Ma [33], [34]. Convolution neural network using multiple distance face [35].

It is also easy to implement as well, such as proposed by Schroff et al. [36], which used triplet loss embedding proposed as a feature that projected face feature into one single point. Our research uses this method because this system has high accuracy while recognizing a face.

Facial recognition was implemented in many attendance systems. For example, in research by Sharanya et al. [37], they use the haar cascade method to detect face and face recognition by the Local Binary Pattern (LBP). This method is not good enough to detect and recognize a face in the attendance system. Another method using deep learning using arcFace was carried out by Son et al. [38]. The results were quite satisfying but still lacked in the response time, which was not so fast. The implementation of face recognition using face recognition was also carried out by Zhi-heng et al. [39] by using a multi-layer perceptron deep learning to recognize a face, but in this study, Zhi-heng encountered a problem

when the number of faces recognized in one frame. Another implementation of face recognition using Eigen Face Recognition [40], [41], but this method has low accuracy. Another approach method using triplet loss embedding was conducted by Schroff et al. [36], namely FaceNet. They use triplet loss embedding value to calculate positive pair and negative pair value, to determine one pair of images are some person or a different person. This method has high accuracy, simple architecture, low computation resource and can be implemented for real-time face recognition. Based on the advantages, we propose a simple, fast, accurate, and robust and can be run on the low-end computer framework of the attendance system with face recognition feature using a triplet loss embedding feature in FaceNet [36] this research.

### B. Research Method

To make a real-time attendance system using triplet loss embedding face recognition, based on FaceNet architecture by Schroff et al. [36] with a fast and accurate authentication process, we adapted code from the OpenFace program using Python programming language, Dlib, and Keras library in the implementation. For the model, we use a pre-trained nn4 small model; a model trained using Keras Library.

In the real-time face recognition process, the process contains two-stage, the first stage is the face detection process, and the second stage is the face classification process. Face image was captured in the first stage and then was classified in the second stage. For the face detection process, we use the Haar Cascade Face detection method by Viola and Jones. This method is used because, in the process of authentication, our system does not require the ability to detect and capture multiple faces at one time, only one face one time, and the user was also cooperative during authentication. And also, we only need a captured face image in a frontal condition up to a 45-degree slope. Haar Cascade method is fair enough for this.

For the reference dataset, we collected several photos from each identity with variations in poses and expressions. We limit the poses not to exceed 45 degrees due to method limitations. After collection, we divided the database into two-part for data train and data test, with a ratio of 5:1. The dataset contains 7550 images from 94 identities, with an average of 80 images for each identity.

Our framework process begins with the process of reading each face image from the reference dataset. The next step is face detection. After the face was detected, the process continued with localization of inner eyes and bottom lips, outer eyes, and nose as a reference point in the alignment process. The alignment process uses facial landmarks based on cascade regressors [42] in the Dlib library; the face image is transposed gradually to the frontal position. Each image before the aligned process bounded by a box and cropped resized to 96 x 96 pixels to make the final result of triplet loss embedding has a standard, and this standard is also used in the process of capturing images in real-time. The same process is applied for an image in real-time face recognition, face captured, and aligned, transpose into frontal position, bounded by box, and resize into 96 x 96 pixels. This same process aims to make both images from the reference dataset and captured face from camera have the same standard, which indirectly increases the accuracy and robustness of the created attendance recording system framework.

To get triplet loss embedding, FaceNet [36] uses a deep convolutional base on Zeiler and Fergus [43] a style network, and recent inception [44]. FaceNet employs triplet loss which directly reflects the task face verification, recognition, and clustering. The architecture strives an embedding  $f(x)$ , from an image  $x$  into a feature space  $\mathbb{R}^d$ , which is the distance between all face images, independent image condition, if the same identity squared distance is small, squared distance between pair of the image from different identity is large. Using a pair of positive and negative used in [45] in eq (1).

Triplet loss from [45] is suitable for face verification; the loss encourages all face one identity to be projected onto a single point in the embedding space, the loss tries to margin between each pair of faces from all other faces, one identity projected into one manifold, and still enforcing the distances to other identities. Triplet loss embedding represent by  $f(x) \in \mathbb{R}^d$ . Its embedded image  $x$  into d-dimensional Euclidean space. Embedding to live on d-dimensional hypersphere, i.e.  $\|f(x)\|_2 = 1$ . This loss motivated to nearest-neighbor classification, ensure image  $x_i^a$  (anchor) of a specific person is closer to all other images  $x_i^p$  (positive) other same person, and far for any image  $x_i^n$  (negative)

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2 \quad (1)$$

$$\forall (f(x_i^a), f(x_i^p), f(x_i^n)) \in \mathcal{T} \quad (2)$$

Where is  $\alpha$  is a margin that enforces between positive and negative pair,  $\mathcal{T}$  is the set all possibility triplets in the training set and has cardinality N. The loss minimized by:

$$\sum \left[ \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right] \quad (3)$$

To produce all possible triplets that can be easily fulfilled by using the constrain in ea. (1). To ensure fast convergence it is crucial to select triplets that maximize the triplet constraint in ea. (1). It means with the given  $x_i^a$ , architecture want to select  $x_i^p$  (hard positive) using  $\argmax_{x_i^p} \|f(x_i^a) - f(x_i^p)\|_2^2$  and similarity  $x_i^n$  (hard negative) with  $\argmin_{x_i^n} \|f(x_i^a) - f(x_i^n)\|_2^2$ .

FaceNet directly learns a mapping from face images to a compact Euclidean space where distance directly corresponds to a measure of face similarity. When space was produced, the tasks face recognition, verification, and clustering that can be easily implemented using the standard technique with embedding feature vectors. FaceNet using triplets of roughly aligned matching/nonmatching face patches was generated using an online triplet mining method based on LMNN [46]. This method learns a Euclidean embedding per image using a deep convolutional neural network, was trained to get squared L2 distance in embedding space directly corresponding with face similarity. The embedding FaceNet has great representational efficiency using only 128 bytes of representational per face. Faces of the same person have a small distance and face different person have large distances. After embedding created the task of face verification simply involves thresholding the distance between the two embeddings; in original FaceNet recognition becomes k-NN classifications problem, and clustering can be achieved by using k-means or agglomerative clustering.

After an alignment process was calculated, each image resulted in triplet loss embedding, which this value calculated for embedding vectors for each image using the pre-trained model used. Use that embedding vector L2 distance to calculate and store onto the L2 distance matrix and the positive-negative matrix pair built for each image in the database. From each score in L2 distance, the matrix calculates the F1 score for each distance below the threshold value. In this paper, we use three threshold values from 0.3 to 1.0, with a 0.01 skip. F1 score was interpreted as a weighted average of the precision and recall, F1 score reached the best value at 1, and the worst score at 0. For multi-class and multi-label cases, this is the average of the F1 score of each class with weighting depending on the average parameter. We also calculate the accuracy score for each distance below the threshold value. From all value F1 score matrices, we try to find the max value and get the threshold from that max value. Then we also calculate the accuracy score from all values under that threshold value.

We classified triplet loss by using an SVM classifier when this triplet loss falls into triplet loss embedding space an identity. If triplet loss embedding from the captured image has a small distance compared with somebody triplet loss embedding, we can assume that the captured image has the same identity. This research also evaluates prediction accuracy and uses triplet loss embedding value classified using k-NN classifier and SVM classifier. The reason we use SVM Classifier is this method is very effective in separating the face embedding vector.

### III. RESULT AND DISCUSSION

#### A. Framework

In the proposed system framework, the attendance system uses RFID and combined facial recognition systems as an authentication process. Before the lecture session begins, the student should tap the RFID card on the terminal provided. The terminal reads the student id contained in the card. This id will be sent to the server to validate and get other student information. The attendance system would command the terminal to begin the face recognition process if the server gave a response that the student is valid.

In face recognition, the attendance system asks students to face the camera to capture their facial images. Capturing process, this image can be done in-camera or video model; at this time, the author uses video mode. The face detection process will be conducted in the captured video to detect the presence of a face. If a face is detected, the next process is to bound this face using a box and align the face into a frontal position.

When reading the image database in the system, each image will be projected into embedding spaces. All faces from one identity are projected onto a single point in the embedding space. In this case, all faces in the database will be projected into 94 single points embedding space because we use 94 identities in the database. Triplet loss tries to margin between each pair of faces from all other faces. The same identity is projected into one manifold and enforcing the distance to other identities from a different person. Triplet loss which is represented by  $f(x) \in \mathbb{R}^d$ . Embedded each image  $x$  is represented by 128-dimension Euclidean space. Embedding

to live on the 128-dimensional hypersphere, this loss is motivated to classify using k-nearest or SVM classification, to ensure the image  $x_i^a$  (anchor) of a specific person is closer to all other images  $x_i^p$  (positive) another same person, and farther away for any image  $x_i^n$  (negative) from a different person. The embedding space from the reference dataset will be used as a reference for recognizing some personal identity.

The next process is authentication. Each student who attends a lecture session is required to record their presence by tapping the RFID card in the recording device in the classroom (see Figure 1).

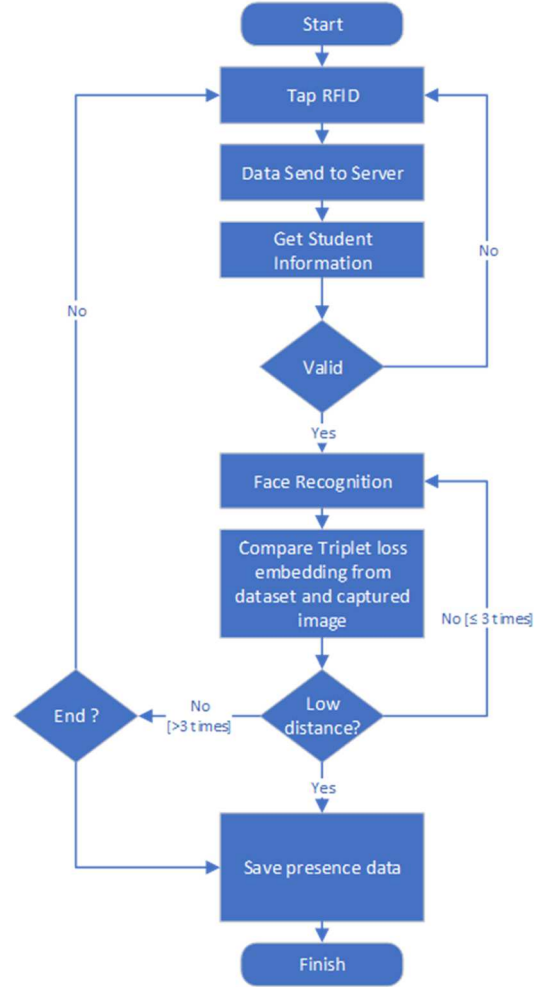


Fig. 1 Authentication Presence Process

After the RFID id card was recorded, the attendance system would read student records from the database as well as the face images. After the student record details are loaded, the process of capturing face images through the camera or video camera begins for 10 seconds. Face images are captured, face image  $y$  presented into 128-dimensional Euclidean space and projected into 128-dimensional hypersphere. A triplet loss calculates from this 128-dimensional Euclidean space. This triplet loss classifies using k-nearest or SVM classification, ensure the image  $x_i^c$  (capture image) of a specific person fall into single point embedding or nothing at all. If image  $x_i^c$  (capture image) fall into a single point embedding. It may be the captured image have identity same with an identity where the image  $x_i^c$  embedding fall. The next phase is to compare the identity process. This authentication process is done with

identity got from the attendance recording phase. If the identity is matched, the authentication process is successful as long as the captured image and images from the reference dataset fall into the same triplet loss embedding vector for the same person. If the face captured image and student's face are predicted as the same person, the system will record the student attendance.

This authentication process will be done over and over again until it is assumed that the image from the database is considered the same as what was caught on the camera, or this check is idle for 10 seconds, or there is no one in front of the camera. This attendance system does not automatically capture images or videos but is triggered by the RFID tapping process. It is hoped that this mechanism will reduce the workload of the system because the request for the face recognition process is only made after the student has tapped and the identity of the RFID is declared to be in the database, so the system does not run all the time to perform facial recognition. Authentication will be declared valid if the identities that appear with the identities obtained from the face recognition process are the same. Next is recording student attendance in the attendance database data; after attendance is recorded, the attendance recording process is complete. For a clear understanding of that process, we showed the flow of how two images compared the 128D classification activity diagram in Figure 1.

### B. Evaluation

In the first phase, we tested the system using face images contained in the dataset and captured images in real-time face recognition. Using the k-NN classification, we get a classification accuracy at  $96.2 \pm 0.1\%$  and using SVM classification at  $95.2 \pm 0.1\%$  as mentioned in table 1.

TABLE I  
ACCURACY COMPARISON CAPTURE IMAGE IN DATASET OUT DATASET

Database	k-NN classifier	SVM Classifier
7550 images from 94 7550	$96.2 \pm 0.1\%$	$95.2 \pm 0.1\%$
7550 images from 94 identity + Captured Images from webcam	$96 \pm 0.2\%$	$95 \pm 0.1\%$

This performance is below baseline OpenFace ( $0.9963 \pm 0.009$  accuracies). The indication is because of the different databases that we used. We used small pre-trained model (nn4 small) which has less accuracy because that model used fewer parameter in which this is good for speeding up our face recognition process. SVM and KNN are used as classifiers because our method uses the triplet loss embedding features, which must be mapped to a vector space. The suitable ones are SVM and KNN.

In the second phase, we tested the system using face images captured from a camera or video camera (we used  $480 \times 640$  pixels resolutions), where the captured image is certainly not in the database. The system almost recognizes everyone whose identity is recorded in the reference dataset. From 100 captured image/video the system can recognize  $95\% \pm 0.2$  accuracies (normal face, without occlusion), 84% accuracy when the head is tilted around 45%, 94% accuracy for small detected faces ( $80 \times 80$  pixels), 75% accuracy with small occlusion on the face, and 20% accuracy for high occlusion.

In the third phase, we tested the system using face images captured from a web camera and adding them into the face image dataset as part of the database (7550 images + 100 images). In this test, we compared the image sample which each embedded in the test. The system was able to recognize almost all photos that were randomly chosen. The performance that happens in the form of accuracy when using face images in a database is  $96\% \pm 0.2$  using k-NN and  $95 \pm 0.1\%$ .

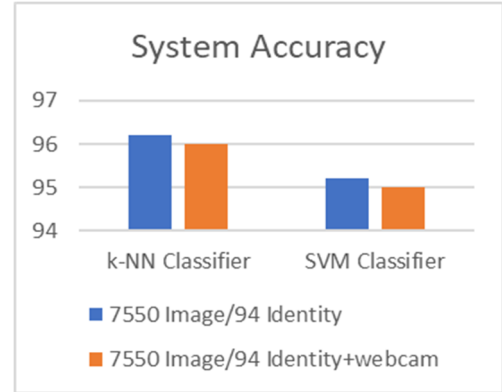


Fig. 2 System Accuracy

### IV. CONCLUSION

Through this experiment, it can be concluded that the face recognition authentication process can be implemented on an attendance recording system, and triplet loss embedding in FaceNet can be used as a robust feature for face recognition in the attendance system because triplet loss has high accuracy. From 100 captured image/video the system can recognize  $95\% \pm 0.2$  accuracies (normal face, in conditions without occlusion), 84% accuracy when the head is tilted around 45%, 94% accuracy for small detected faces ( $80 \times 80$  pixels), 75% accuracy with small occlusion on the face, and 20% accuracy for high occlusion. This method is also suitable to implement in real-time face recognition because the architecture is quite lightweight and easy to implement. Low accuracy in high occlusion is not a problem in this attendance system, because the student can repeat the process until authentication success.

The attendance recording system using RFID and combined facial recognition systems as an authentication process will increase student attendance in lectures. The system is difficult to fake because it uses face recognition as an authentication process. Facial biometric marks use in this attendance system will be very difficult to fake, as a result, students will always be present in lectures, which in turn will improve the quality of the existing education process.

In the future, the result can be improved by using high camera resolution. For a better authentication process, a face recognition system can be added with facial expression recognition. With this feature, users are required to perform an expression instructed by face recognition, using database and YOLO method for better performance.

### ACKNOWLEDGMENT

This research was funded by Bina Nusantara University's applied research grant year 2019.



## REFERENCES

- [1] T. Fadelelmoula, "The impact of class attendance on student performance," *Int. Res. J. Med. Med. Sci.*, vol. 6, no. May, pp. 47–49, 2018.
- [2] G. A. Tetteh, "Effects of Classroom Attendance and Learning Strategies on the Learning Outcome," *J. Int. Educ. Bus.*, vol. 11, no. 2, pp. 195–219, 2018.
- [3] B. Bartanen, "Principal Quality and Student Attendance," *Educ. Res.*, vol. 49, no. 2, pp. 101–113, 2020.
- [4] Y. Zhang *et al.*, "A Teaching Evaluation System Based on Visual Recognition Technology," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 782, no. 3, 2020.
- [5] K. Lee Lerner and Brenda Wilmoth Lerner, *World of forensic science*. 2005.
- [6] S. Z. Li and A. K. Jain, *Introduction Face Recognition*. 2011.
- [7] K. Ye and F. Hu, "Research on Cross-Age Face Verification Based on Artificial Neural Network Under Examination Environment," 2017 10th Int. Symp. Comput. Intell. Des., pp. 430–433, 2017.
- [8] N. Bakshi and V. Prabhu, "Face recognition system for access control using principal component analysis," *ICCT 2017 - Int. Conf. Intell. Commun. Comput. Tech.*, vol. 2018-Janua, pp. 145–150, 2018.
- [9] N. A. Abdullah, M. J. Saidi, N. H. A. Rahman, C. C. Wen, and I. R. A. Hamid, "Face recognition for criminal identification: An implementation of principal component analysis for face recognition," *AIP Conf. Proc.*, vol. 1891, 2017.
- [10] L. Wang and A. A. Siddique, "Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology," *Meas. Control (United Kingdom)*, vol. 53, no. 7–8, pp. 1070–1077, 2020.
- [11] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [12] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Networks*, vol. 13, no. 6, pp. 1450–1464, Nov. 2002.
- [13] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, 1997.
- [14] H. Hu, P. Zhang, and F. De la Torre, "Face recognition using enhanced linear discriminant analysis," *IET Comput. Vis.*, vol. 4, no. 3, p. 195, 2010.
- [15] Jian Yang, D. Zhang, A. F. Frangi, and Jing-yu Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
- [16] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2106–2112, 2010.
- [17] R. Rahim, T. Afriliansyah, H. Winata, D. Nofriansyah, Ratnadewi, and S. Aryza, "Research of Face Recognition with Fisher Linear Discriminant," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 300, no. 1, 2018.
- [18] X. He, S. Yan, Y. Hu, P. Niyogi, and H. J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, 2005.
- [19] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [20] A. Bolotnikova, H. Demirel, and G. Anbarjafari, "Real-time ensemble based face recognition system for NAO humanoid using local binary pattern," *Analog Integr. Circuits Signal Process.*, vol. 92, no. 3, pp. 467–475, 2017.
- [21] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor phase patterns (HGPP): A novel object representation approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 57–68, 2007.
- [22] B. Yang and S. Chen, "A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image," *Neurocomputing*, vol. 120, pp. 365–379, 2013.
- [23] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative common vectors for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 1, pp. 4–13, 2005.
- [24] Y. Wen, "An improved discriminative common vectors and support vector machine based face recognition approach," *Expert Syst. Appl.*, vol. 39, no. 4, pp. 4628–4632, 2012.
- [25] E. Kokiopoulou and Y. Saad, "Orthogonal neighborhood preserving projections: A projection-based dimensionality reduction technique," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2143–2156, 2007.
- [26] M. Z. Al-Dabagh, M. H. Mohammed Alhabib, and F. H. AL-Mukhtar, "Face Recognition System Based on Kernel Discriminant Analysis, K-Nearest Neighbor and Support Vector Machine," *Int. J. Res. Eng.*, vol. 5, no. 2, pp. 335–338, Mar. 2018.
- [27] X. M. Zhao and C. B. Wei, "A real-time face recognition system based on the improved LBPH algorithm," 2017 IEEE 2nd Int. Conf. Signal Image Process. ICSIP 2017, vol. 2017-Janua, pp. 72–76, 2017.
- [28] R. Singh, M. Vatsa, and A. Noore, "Face recognition with disguise and single gallery images," *Image Vis. Comput.*, vol. 27, no. 3, pp. 245–257, 2009.
- [29] E. L.-M. Huang, Gary B., Honglak Lee, "Learning Hierarchical Representations for Face Verification.pdf," pp. 978-1-4673-1228-8/12, 2012.
- [30] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," in 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, vol. 73, no. 6, pp. 1701–1708.
- [31] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 2892–2900, 2015.
- [32] X. Xu, H. A. Le, P. Dou, Y. Wu, and I. A. Kakadiaris, "Evaluation of a 3D-aided pose invariant 2D face recognition system," *IEEE Int. Jt. Conf. Biometrics, IJCB 2017*, vol. 2018-Janua, pp. 446–455, 2018.
- [33] S. Qiao and J. Ma, "A Face Recognition System Based on Convolution Neural Network," *Proc. 2018 Chinese Autom. Congr. CAC 2018*, pp. 1923–1927, 2019.
- [34] M. Nakada, H. Wang, and D. Terzopoulos, "AcFR: Active Face Recognition Using Convolutional Neural Networks," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017-July, pp. 35–40, 2017.
- [35] H. M. Moon, C. H. Seo, and S. B. Pan, "A face recognition system based on convolution neural network using multiple distance face," *Soft Comput.*, vol. 21, no. 17, pp. 4995–5002, 2017.
- [36] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 815–823, 2015.
- [37] Sharanya T, "Online Attendance using Facial Recognition," *Int. J. Eng. Res.*, vol. V9, no. 06, pp. 202–207, 2020.
- [38] N. T. Son *et al.*, "Implementing CCTV-based attendance taking support system using deep face recognition: A case study at FPT polytechnic college," *Symmetry (Basel)*, vol. 12, no. 2, 2020.
- [39] L. Zhi-heng and L. Yong-zhen, "Design and Implementation of Classroom Attendance System Based on Video Face Recognition," 2019 Int. Conf. Intell. Transp. Big Data Smart City, pp. 385–388, 2019.
- [40] E. Rekha and P. Ramaprasad, "An efficient automated attendance management system based on Eigen Face recognition," in 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017, vol. 5, pp. 605–608.
- [41] N. K. Ayu Wirdiani, T. Lattifia, I. K. Supadma, B. J. Kemanang Mahar, D. A. Nadia Taradhita, and A. Fahmi, "Real-Time Face Recognition with Eigenface Method," *Int. J. Image, Graph. Signal Process.*, vol. 11, no. 11, pp. 1–9, Nov. 2019.
- [42] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1867–1874, 2014.
- [43] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8689 LNCS, no. PART 1, pp. 818–833, 2014.
- [44] C. Szegedy *et al.*, "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 1–9, 2015.
- [45] Y. Sun, X. Wang, and X. Tang, "Deep Learning Face Representation by Joint Identification-Verification," pp. 1–9, 2014.
- [46] K. Q. Weinberger and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," *J. Mach. Learn. Res.*, vol. 10, pp. 173–179, 2007.