Smart Attendance System based on Improved Facial Recognition

Thai-Viet Dang [1,*]
[1] Department of Mechatronics Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam
Email: [1] viet.dangthai@hust.edu.vn
*Corresponding Author

***Abstract*—Nowadays, the fourth industrial revolution has achieved significant advancement in high technology, in which artificial intelligence has had vigorous development. In practice, facial recognition is one most essential tasks in the field of computer vision with various potential applications from security and attendance system to intelligent services. In this paper, we propose an efficient deep learning approach to facial recognition. The paper utilizes the architecture of improved FaceNet model based on MobileNetV2 backbone with SSD subsection. The improved architecture uses depth-wise separable convolution to reduce the model size and computational volume and achieve high accuracy and processing speed. To solve the problem of identifying a person entering and exiting an area and integrating on advanced mobile devices limits to (such as limited memory and on-device storage) highly mobile resources. Especially, our approach yields better results in practical application with more than 95% accuracy on a small dataset of the original face images. Obtained frame rate (25 FPS) is very favorable compared to the field of facial recognition using neural network. Besides, the deep learning based on solution could be applicable in many low-capacity hardware or optimize system's resource. Finally, the smart automated attendance systems is successfully designed basing on the improved efficient facial recognition.**

Keywords—Artificial intelligence; Attendance system; Facial recognition; Internet of Thing (IoT); MobileNets

I. INTRODUCTION

Nowadays, the fourth industrial revolution has achieved a incredible convergence of emerging technological breakthroughs, which is an opportunity to solve the challenges that exist worldwide. In a wide range of real life and practical systems, attendance plays a pivotal role in determining academic performance of learners [1-9]. Almost automatic human identification systems are based on traditional methods such as fingerprints, passwords, and ID's users [1, 2]. But forgetting a password and losing an ID card becomes inconvenient problems. Hence, new technologies have been applied into detecting participation, such as such as the use of QR Code [3, 4] or RFID [5, 6]. QR Codes cannot be read by linear barcode scanners. Therefore, they require camera-based image processing systems in fixed reading distance. Sultana et. al., in [7] presented a attendance tracking system based on Android by extracting the GPS data. Then, Mahesh et. al. in [8] integrated a smartphone with smart classroom based on facial recognition technology to improve the quality of monitoring. RFID depends on the physical frequency of use. The quality is affected by obstacles or wave-absorbing materials near the RFID reader. Other solutions introduce the use of biometrics' indicators to acquire attendance including as follows: voice, fingerprints and face recognition [9-14]. In [10-12], fingerprints have also been used to identify and calculate the attendance number based on the time duration reports. Face recognition has many limitations in a real-world scenario because of lighting conditions and low-quality image processing [9, 13, 14]. Improved face recognition algorithm combines with support systems and devices ensure the stability and effectiveness of the attendance quality in the working environment conditions [13].

Outstanding the face recognition problem [13-18] that requires high accuracy and good processing speed, research work for practical applications, etc. Despite recent advances in facial recognition, the large-scale implementation of effective facial recognition and verification presents severe limitations to current approaches. How to determine if the current image is correct with the information and whether the person's face is in the system or not by comparing the pre-selected facial features from the image database. Pawar et al. in [15] used Local Binary Pattern (LBP) descriptors for converting the input image to a binary image. Then, the descriptors were extracted at multiple resolutions and concatenated into a regional feature vector. Continuously, the calculation of the histogram density on each block gave the histogram feature. Unfortunately, the feature extraction from histograms may be affected by external factors such as input image quality, light, etc. In order to minimize this influence, the article applied FaceNet for calculating the distance between face vectors. Shebani et. al. in [16] proposed the modified architecture of facial recognition including three patch LBP and four patch LBP were combined with Linear Discriminant Analysis (LDA) [16] and Support Vector Machine (SVM) [18-21] to perform face recognition accuracy by encoding similarities between neighboring patches of pixels. Among them, we consider about the problems with limited training data. Especially, the performance of facial recognition system need to be optimized with required accuracy and frame rate under limitation of system resources.

Indeed, previous papers using Neural Network (NN) combined with LDA [22-29] have achieved facial recognition accuracy of more than 95% and processing speed of 4 FPS but due to many multi-layer's perceptron (is a network with n ($n \geq 2$) layers (usually the input is not taken into account): where there is an output layer (nth layer) and (n-1) hidden layer) which increases the computational volume. Pawar et al., in [15] proposed real-time face recognition with LPB model applying into smart city model with only achieves 80 % accuracy [24]. Then, T. V. Dang et. al. improved the facial recognition model in [28] to increase the accuracy of 87-90 %. Besides the feature extraction from histogram has affected the recognition process. Hence, Convolutional Neural Networks (CNNs) have been developed as powerful models for image recognition problems requiring large-scale labeled training data. However, estimating millions parameters of deep CNNs requires a huge amount of labeled samples, restricting CNNs being applied to problems with limited training data [25]. Due to the successful application of deep learning, Deep Convolutional Neural Networks (DCNNs) improves the of face recognition [26]. MobileNetV2 is one of the most popular architectures of DCNNs in computer vision [26, 27]. Furthermore, the MobileNetV2 will optimize face recognition process by reducing the load on the convolutional network layers and replacing the Conv2D layer with a depth-wise separable convolution layer. In addition, Single-Shot multibox Detection (SSD) uses the MobilenetV2's output as the input for face detection. For the purposes of this study, the paper utilizes improved FaceNet model based on MobileNetV2 backbone with SSD subsection, which ensures solving the problem of object identification and can be applied to security or attendance systems using less-resourced mobile devices with high obtained accuracy and speed. The output shows that this is an optimal model for low-resourced mobile and embedded devices as follows: Jetson Nano (quad-core ARM Cortex-A57 CPU, 128-core Maxwell™ GPU) [29, 50-64]. According to practical examples, obtained face images reach to good detection performance with an accuracy of 95% and a fast inference speed of 25 FPS. For small datasets and less resources in training model, the model is more efficient and faster than the state-of-the-art models which require larger datasets for training and processing. Finally, the smart attendance system is successfully designed based on the improved facial recognition and IoT technology [30-36].

The rest of the paper is organized as follows. In Section 2, the author describes the system design and related works in facial recognition of FaceNet model based on MobileNetV2 backbone with SSD subsection. In Section 3, the paper briefly introduces our proposed experimental system. Next, practical experiments of smart attendance system are conducted to compare with previous methods. Finally, a conclusion is presented in Section 4.

II. PROPOSED METHOD

A. System Design

The structure of the automated attendance system is divided into two main processing tiers as shown in Fig. 1. At the user interface layer, there will be two layers of identification for optimal accuracy: identification and attendance based on fingerprint system and second attendance with camera system. After the data is collected at the first layer, transferring to the second layer for data processing. At here, the data will be completely calculated. Next, the lecturers will evaluate the final student's score and save the data to the database at the end of the course. Finally, the data can be imported and exported by excel file.

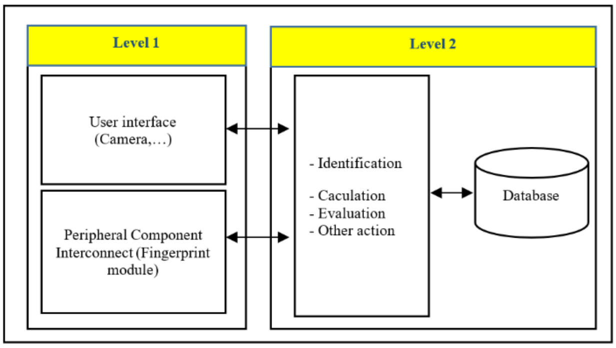


Fig. 1. The detailed structure of attendance monitoring system

Hardware:

Jetson Nano in Fig. 2: Jetson Nano 4GB embedded computer can run the model with low power from 5-10 Watt. Jetson Nano 4GB developer AI Artificial Intelligence Development Board including as follows: GPU: 128-core Maxwell; CPU: Quad-core ARM A57@1.43GHZ; Memory: 4 GB 64-bit LPDDR4 25.6 GB/s and Storage: 16GB EMMC. With the features that Jetson Nano, especially the GPU-based board offers, We can deploy a practical attendance system with multiple recognition cameras requiring cameras with good enough resolution and from 25 FPS to 30 FPS.

Fig. 2. Jetson nano with 8MP IMX219

Fingerprint scanning module R305 in Fig. 3: a product with a small size, stable performance, and simple structure. Specifications: power supply from 3.6-6 VDC, communication type is USB 1.1 or TTL-UART, multi-fingerprint recognition mode, baud rate: 57600 bps, average recognition time is less than 0.8s, etc. It helps scan and saves the user's fingerprint identification information, distinguishing it from fingerprints. Therefore, we use the fingerprint scanning module combined with Jetson Nano to enhance the attendance system and check identification functions.

Fig. 3. Fingerprint scanning module R305.

Software:

Use PYQT5, a utility tool for designing graphical user interfaces (GUIs). This is an interface written in python programming language for Qt, one of the most popular and influential cross-platform GUI libraries. Then, use some deep learning frameworks: Keras, Tensorflow, etc., for learning models [30].

B. Data Processing

With the attendance system using a normal scan module, the whole data system will be processed on a single thread. However, a system combining fingerprints with facial recognition, the hardware system is depicted block diagram as shown in Fig. 4. Face recognition system using Jetson Nano capable of handling multiple video streams [31]. Jetson Nano is connected to the camera module for storing the acquired camera images. An observation device that connects to the Jetson Nano such as a TV or LCD monitor. Jetson Nano also supports Internet access with smart devices: mobile phones, laptops, and desktop computers [31-33].

Fig. 4. Block diagram of the proposed hardware system

Based on the proposed practical system, Fig. 5 shows the process of identifying and processing data.

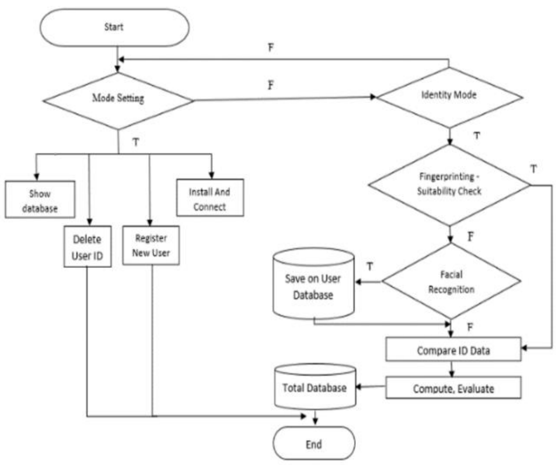
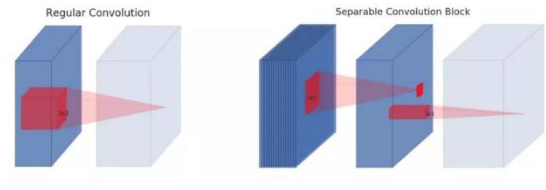


Fig. 5. The flow chart of data processing

C. MobileNetV2 backbone and SSD subsection

MobileNetV2 backbone:

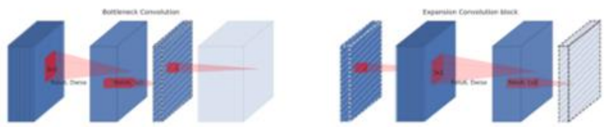
MobileNets are based on a streamlined architecture that uses depth-wise separable convolutions to build light weight deep neural networks in Fig. 6 [37-44].



(a) (b)

Fig. 6. Depthwise Separable Convolutions with (a) regular and (b) separable

MobileNetV2 [37-43] improves over MobileNetV1 [26, 44] to gain higher accuracy with fewer parameters and fewer calculations. In this section, we mainly introduce the core features of the MobileNetV2 to be used, the optimization of the loss function and utilizes the improved model architecture from FaceNet model. Fig. 7 illustrates MobileNetV2 using Depth-wise Separable Convolutions with linear bottlenecks and inverted residual block (shortcut connections between bottlenecks) [43].



(a) (b)

Fig. 7. Separable convolutional blocks with (a) Separable with linear bottleneck and (b) bottleneck with passion layer.

Subsequently, MobileNetV2's residual block is the opposite of traditional residual architectures, because the traditional residual architecture has a larger number of channels at the input and output of a block than the intermediate layers. Among layers, there is an inverted residual block, we also use depth-separated convolution transforms to minimize the number of model parameters. The solution helps the MobileNet model to be slightly reduced in size. The detailed structure of a depth-separated convolution of bottleneck with residuals is as follows in Fig. 8.

Input	Operator	Output
$h \times w \times k$	1x1 conv2d , ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3x3 dwise s=s, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1x1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

Fig. 8. Bottleneck residual block converts from k to channel k' with stride and expansion factors.

In the era of mobile networks, the need for lightweight and real-time networking is growing. However, many identity networks cannot meet the real-time requirements due to too many parameters and computations. To solve this problem, the proposed method using the MobileNetV2 backbone achieves superior performance compared with other modern methods in the database of facial expressions and features in Fig. 9.

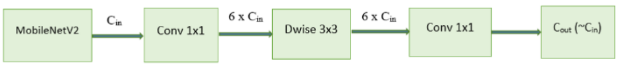


Fig. 9. Infrastructure of MobileNetV2 network

Fig. 9 presents MobileNetV2's workflow using 1x1 point convolution to extend the input channels. Then use depth convolution for input linear feature extraction and linear convolutional integration to combine output features while reducing the network size. After size reduction, replace Relu6 with a linear function to enable the output channel size to match the input. MobileNetV2 will be very useful when applied to SSD to reduce latency and improve processing speed. Besides Fig. 11 depicts subsection SSD [45, 46] which is one of the fast and efficient object detection features with better minimal processing time than YOLO [14] and Faster-RCNN [46]. Consequently, the subsection SSD uses MobilenetV2 backbone to extract the feature map and adds extra bits to predict the object (see Fig. 10).

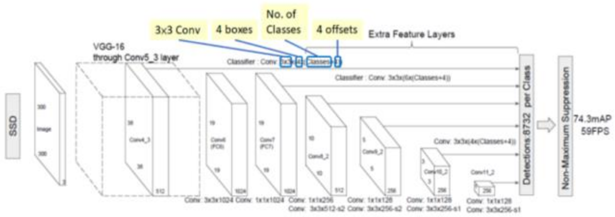


Fig. 10. The architecture of SSD network.

D. FaceNet Embeddings

FaceNet uses initiation modules in blocks to reduce the number of trainable parameters [13]. This model takes a 160×160 RGB image and generates a 128-d embedding vector for each image. FaceNet features extraction for face recognition. Use Facenet to decompose facial features into vectors and then use the triplet loss function to calculate the distance between face vectors.

First, the face will be represented as a vector so that the mathematical comparison and recognition process becomes easy. In face identification, we will need to calculate the similarity and difference between the faces we get, and in essence we will refer to the problem of calculating the distance between the vectors of Triplet Loss. Triplet is a set of three parameters we include: one face image of any person (query), one other face image of that person (positive), one different person's face image (negative) in Fig. 11.

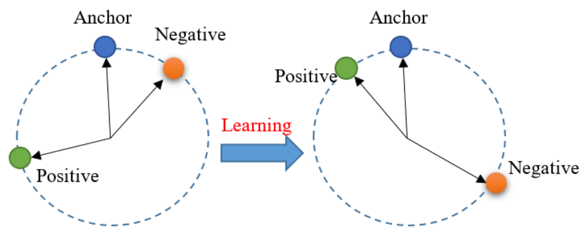


Fig. 11. Triplet Loss

Second, instead of using conventional loss functions that only compare the output value of the network with the actual ground truth of the data, Triplet Loss introduces a new formula that includes three input values, including the output anchor of the network, positive x_i^p: the photo is the same person with the anchor, positive x_i^p: the photo is not the same person with the anchor.

$$\left\|f(x_i^a) - f(x_i^b)\right\|_2^2 + \alpha\forall\left(f(x_i^a), f(x_i^b), f(x_i^n)\right) \in T < \left\|f(x_i^a) - f(x_i^n)\right\|_2^2 \quad (1)$$

where α is the distance between the positive and negative pair, the minimum necessary deviation between the two values, $f(x_i^a)$ is the embedding of x_i^a. The above formula shows that the distance between two embeddings $f(x_i^a)$ and $f(x_i^p)$ will have to be at least α-value less than the pair $f(x_i^a)$ and $f(x_i^n)$. In other words, after training, the result obtained is that the difference between the two sides of the formula is as large as possible (meaning x_i^a get close to x_i^p).

Finally, a pre-training image database is taken from the Labeled Face in Wild (LFW) dataset consisting of more than 13000 images containing human faces [28, 47, 48] to obtain the trained FaceNet model with the architecture in Fig. 12.

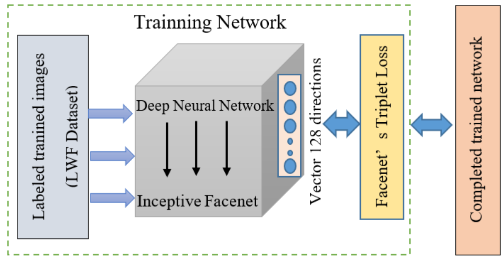


Fig. 12. The architechture of FaceNet model.

E. Facial Recognition

The recognition process consists of two main processes, face detection and face recognition. Each process uses different algorithms. In this paper, the author uses multilayer convolutional neural network to detect faces in frames and FaceNet feature extraction for face recognition. Below is the image data processing cycle in Fig. 13. During the face recognition process, the subsection SSD uses the MobileNetV2 backbone network to define the face's box. But the last few layers of the network, such as FC, Maxpool, and Softmax will be ignored, and the output of MobileNetV2 will be used as feature maps to base the detection from the input images.

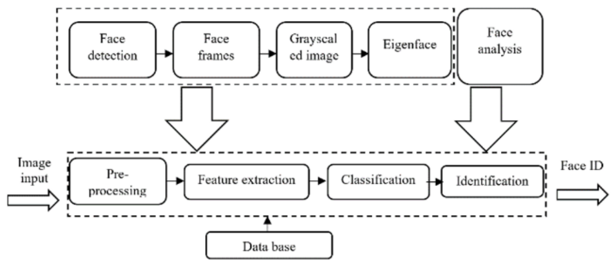


Fig. 13. Facial recognition process.

Then, FaceNet takes an image of each person's face as input and outputs a vector of 128 numbers representing the most important facial features. In machine learning, this vector is called embedding. Next, is a classifier used to get the distance of the facial features to distinguish between several identities. Some of the most widely used algorithms in face recognition tasks are Support Vector Machine (SVM) in [18, 19], and K-Nearest Neighbor (KNN) in [50] because of their efficiency in multi-class classification. With face detection in frame, FaceNet model extract feature for face recognition at the final step in Fig. 14.

Fig. 14. FaceNet model based on MobileNetV2 backbone and SSD subsection

III. Results and Discussion

A. Experimental results

First: With a multi-layer convolutional network, the pre-training image database is taken from the Labeled Face in Wild (LFW) dataset consisting of more than 13000 images containing human faces [28, 47, 48, 49] in Fig. 15.

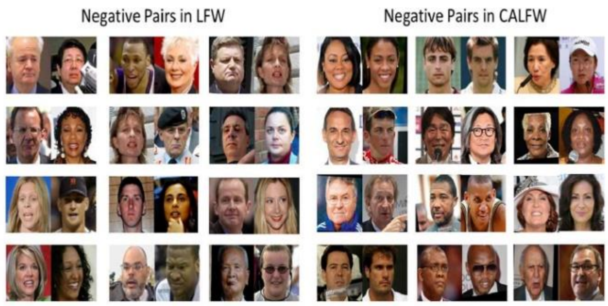


Fig. 15. Labelled Face in Wild (LFW) Dataset

Second: The attendance training data is a set of images of 150 labeled people from three classrooms. Each classroom has 50 students will 50 their images to train for face recognition. Most of the faces of the dataset appearing in the direct face view image with full of feature information (see Fig. 16).

Fig. 16. The sample of image for attendance training data

We evaluate the FaceNet model based on the MobileNetV2 backbone network and SSD subsection using in a broader face dataset and compare it with models such as MTCNN using O-net, P-net, R-net in [7, 43, 48]. Besides, the test results in Fig. 16 show that the SSD-MobileNetV2 achieves outstanding performance compared to the other models on the Jetson Nano embedded computer. First, Fig. 17(a) only uses MTCNN. Second, Figs. 17(b) and 17(c) use Rectinaface model based on the backbone R512 and R50, respectively [47]. All of models in [7, 43, 48] gain the low frame rate at the same hight accuracy. Similarly, to the result of Rectinaface model based on MobilenetV2. Specially, the FaceNet model based on the backbone MTCNN [28] using for attendance system is about 87 to 90 %. Finally, Fig. 17(e) shows the results of detected faces with SSD-MobileNetV2 obtaining an accuracy of 97-99 % and a frame rate of 20-25 FPS.

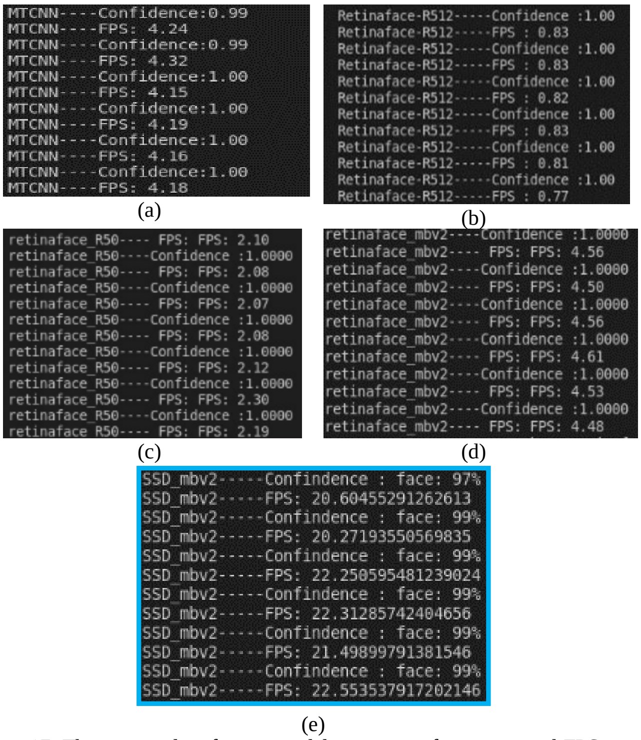


Fig. 17. The test results of some models in terms of accuracy and FPS rate with (a) MTCNN, (b) Retinaface-R512, (c) Retinaface-R50, (d) Retinaface-MobileNetV2, and (e) SSD-MobileNetV2, respectively.

The comparison among above facial recognition models is shown more clearly in Table 1.

TABLE 1. TEST RESULTS OF EXPERIMENTAL FACE RECOGNITION MODELS.

Model	Backbone	Dataset	Accuracy (%)	FPS
MTCNN	O-net, P-net, R-net	widerface	from 97 to 100	4.15 - 4.32
RectinaFace	Renet-512	widerface	100	0.77 - 0.83
RectinaFace	Resnet-50	widerface	100	2.07 - 2.30
RectinaFace	MobileNetV2	widerface	100	4.48 - 4.61
SSD	**MobileNetV2**	**widerface**	**from 97 to 99**	**20.25 - 22.55**

We evaluate the performance of FaceNet model based on MobilenetV2-SSD in face recognition algorithm with 14 random students' faces. The model achieves the accuracy of from 91% to 94% as shown in Fig. 18. The experimental results have been really better than FaceNet based on the backbone MTCNN in [28].

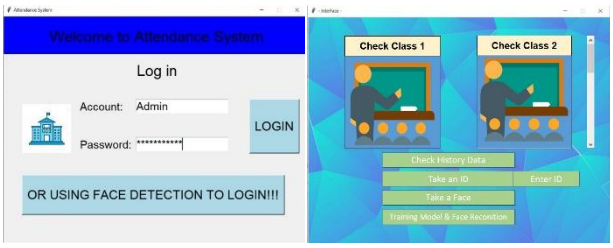
Fig. 18. Face recognition results based on MobileNetV2 backbone with SSD subsection

B. *Smart Attendance system*

The system operation is according to step by step-in real-time activities as follows.

Step 1: Login system in the first time by password and account of admin in Fig. 19(a). Then, use face recognition function to login later after identifying and training successfully.

Step 2: Update the user's face image in the main interface. All steps of facial recognition in Fig. 19(b) will be completely performed.



(a) (b)

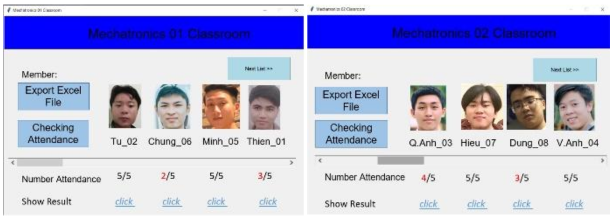
Fig.19. The system interface with (a) login and (b) main interface of attendance system.

Step 3: Collect data of student face's image to train the face recognition model, next the system will automatically calculate the attendance score based on the detecting student's participation frequency in Fig. 20.

(a) (b)

Fig. 20. The attendance system interface with (a) Class 1 and (b) Class 2

Step 4: Export or import data by a excel file or send a email to admin. The lecturer also monitors student's diligence in Fig. 21 each classroom, respectively.



(a) (b)

Fig. 21. The attendance system interface with (a) Class 1 and (b) Class 2.

Step 5: Evaluate students by aggregate data in Fig. 22.

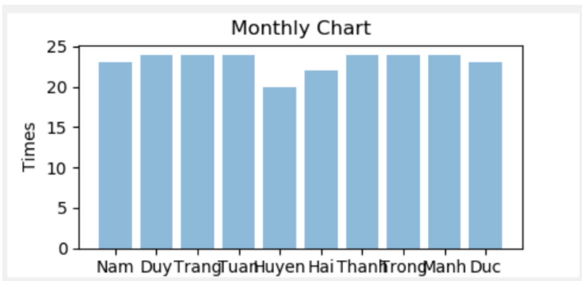


Fig. 22. The monthly chart of student's diligence process.

Based on the proposed experimental system, for small datasets and less resources in training model, the updated

FaceNet model is more efficient and faster than the previous state-of-the-art models which require larger datasets for training and processing.

In terms of the ability of the system improving authenticity of the attendance, the purposed system illustrated two important attributes as follows. First, use automatic authentication of attendance, by face recognition, because it is the least evasive and require basic acquisition devices. Second, use small datasets and less resources in training model, and integrate on advanced limited mobile devices (such as limited memory and on-device storage) to highly mobile devices resource.

IV. CONCLUSIONS

This paper utilizes a FaceNet model based on the backbone MobilenetV2 with the subsection SSD to identify faces using depth-separated convolutional networks to reduce the model size and computational volume, with awe-inspiring results. The authors experimented with and evaluated the proposed model based on comparing different Rectinal models and MCTT backbone. Using the same dataset as widerface, SSD combined with MobileNetV2 backbone has ensured accuracy of about 99% in simulated experiments and 91-95% in practical applications. The speed of process is effectively improved with the frame rate of 20-23 (FPS). For small datasets and less resources in training model, the updated FaceNet model is more efficient and faster than the previous state-of-the-art models which require larger datasets for training and processing. Besides, the deep learning based on solution could be applicable in many low-capacity hardware or optimize system's resource. To improve the accuracy and speed rate, the input data is continuously refined and updated based on the peripheral factors. Hence, the author's next research direction is not only the promotions in system's precisions and speed, but also the security and attendance enhancement with 3D facial recognition. Moreover, it will strengthen anti-face spoofing methods to increase system security. Finally, obtained results can apply for intelligent IoT system and smart services.

CONFLICT OF INTEREST

The author declares no conflict of interest

ACKNOWLEDGMENT

This research is funded by Hanoi University of Science and Technology (HUST) under project number T2022-PC-029.

REFERENCES

[1] G. Hua, M. H. Yang, E. L. Miller, T. M. Ma, D. J. Kriegman, and T. S. Huang, "Introduction to the special section on real world face recognition," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 33, pp. 1921-1924, 2011.

[2] F. P. Filippidou, and G. A. Papakostas, "Single Sample Face Recognition Using Convolutional Neural Networks for Automated Attendance Systems," *2020 Fourth Int. Conf. Intell. Comput. Data Sci. (ICDS)*, 2020.

[3] M. G. M. Johar, and M. H. Alkawaz, "Student's activity management system using QR code and C4. 5 algorithm," *Int. J. Med. Toxicology & Legal Med.*, vol. 21, pp. 105-107, 2018.

[4] F. Masalha, and N. Hirzallah, "A students attendance system using QR code," *Int. J. Adv. Comput. Sci. Appl.*, vol. 5, pp. 75-79, 2014.

[5] O. Arulogun, A. Olatunbosun, O. Fakolujo, and O. Olaniyi, "RFID based students attendance management system," *Int. J. Sci. & Eng. Res.*, vol. 4, pp. 1-9, 2013.

[6] F. Silva, V. Filipe, and A. Pereira, "Automatic control of students' attendance in classrooms using RFID," *ICSNC 08. 3rd Int. Conf. Syst. Netw.*, pp. 384-389, 2008.

[7] M. Karunakar, C. A. Sai, K. Chandra, and K. A. Kumar, "Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environment," *Int. J. Recent Develop. Sci. Technol.*, vol. 4, no. 5, pp. 194-201, 2020.

[8] S. Wenhui, and J. Mingyan, "Face Recognition Based on Multi-view - Ensemble Learning," *Chin. Conf. Pat. Recognit. Comput. Vision (PRCV)*, pp. 127-136, 2018.

[9] A. S. Al-Waisy, R. Qahwaji, S. Ipson, and S. Al-Fahdawi, "A Robust Face Recognition System Based on Curvelet and Fractal Dimension Transforms," *IEEE Int. Conf. Comp. Inf. Technol.*, pp. 548-555, 2015.

[10] A. Fassio et al., "Prioritizing Virtual Screening with Interpretable Interaction Fingerprints," *J. Chem. Inf. Model.*, vol. 62, no. 18, pp. 1-53, 2022.

[11] Y. Jiang, "Space Debris Fingerprint Definition and Identification," *The Journal of Brief Ideas*, 2022.

[12] A. B. V. Wyzykowski, M. P.Segundo, and R. D. P. Lemes "Multiresolution synthetic fingerprint generation," *IET Biometrics*, vol. 11, no. 1, 2022.

[13] W. Chunming, and Z. Ying, "MTCNN and FaceNet Based Access Control System for Face Detection and Recognition," *Autom. Control Comput. Sci.*, vol. 55, pp. 102-112, 2021.

[14] A. Bochkovskiy, C.Y. Wang, and H. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934.*, 2020.

[15] S. Pawar, V. Kithani, S. Ahuja, and S. Sahu, "Local Binary Patterns and Its Application to Facial Image Analysis," *2011 Int. Conf. Recent Trends Inf. Technol. (ICRTIT)*, pp. 782-786, 2011.

[16] Q. A. Shebani, "A Hybrid Feature Extraction Technique for Face Recognition," *Int. Proc. Comput. Sci. Inf. Technol.*, vol 3, no. 2, 2012.

[17] A. Sharma and S. Chhabra, "A Hybrid Feature Extraction Technique for Face Recognition," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 7, no. 5, pp. 341-350, 2017.

[18] M. Mia, R. Islam, M. F. Wahid, and S. Biswas, "Image Reconstruction Using Pixel Wise Support Vector Machine SVM Classification," *Comput. Sci. Math. Int. J. Sci. Technol. Res.*, vol. 4, no. 2, pp. 232-235, 2015.

[19] U. Maulik and D. Chakraborty, "Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 1, pp. 33-52, 2017.

[20] T. V. Dang, "Smart home Management System with Face Recognition based on ArcFace model in Deep Convolutional Neural Network," *J. Robot. Control (JRC)*, vol. 3, no. 6, pp. 754-761, 2022.

[21] T. V. Dang et at., "Design of a Face Recognition Technique based MTCNN and ArcFace," MMMS 2022, LNME, 2022.

[22] Z. Haitao, L. Zhihui, L. Henry and Z. Xianyi, "Linear Discriminant Analysis," *Feature Learn. Understanding*, pp 71-85, 2020.

[23] F. Zuo, and P. H. N. de With, "Real-time Face Recognition for Smart Home Applications," *2005 Digest of Tech. Papers, Int. Conf. Consum. Electron.*, pp. 35-36, 2005.

[24] S. Pawar, V. Kithani, S. Ahuja, and S. Sahu, "Smart Home Security using IoT and Face Recognition," *2018 Fourth Int. Conf. Comput. Commun. Control and Automat. (ICCUBEA)*, vol. 176, no. 13, pp. 45-47, 2018.

[25] F. P. Filippidou, and G. A. Papakostas, "Single Sample Face Recognition Using Convolutional Neural Networks for Automated Attendance Systems," *2020 Fourth Int. Conf. Intell. Comput. Data Sci. (ICDS)*, 2020.

[26] A. G. Howard, M. Zhu, Bo Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *Comput. Vision Pat. Recognit.*, 2017.

[27] A. Wirdiani, P. Hridayami, and A. Widiari, "Face Identification Based on K-Nearest Neighbor," *Sci. J. Inform.*, vol. 6, no. 2, pp. 150-159, 2019.

[28] T. V. Dang, and D. K. Nguyen, "Research and Design the Intelligent Mechatronics system applying with face recognition and deep learning in student's diligencet," *7th Nat. Conf. Sci. Eng. Meas.*, pp. 239-246, 2020.

[29] A. Süzen, B. Duman, and B. Sen, "Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN," *2020 Int. Congr. Human-Comput. Interact. Optim. Robot. Appl. (HORA)*, pp. 1-5, 2020.

[30] S. Kurzadkar, "Hotel management system using Python Tkinler GUI," *Int. J. Comput. Sci. Mobile Comput.*, vol. 11, no. 1, pp. 204-208, 2022.

[31] V. Sati, S.M. Sánchez, N. Shoeibi, A. Arora, and J.M. Corchado, "Face Detection and Recognition, Face Emotion Recognition Through NVIDIA Jetson Nano," *11th Int. Symp. Ambient Intell.*, pp. 177-185, 2020.

[32] M. H. Widianto, A. Sinaga, and M. A. Ginting, "A Systematic Review of LPWAN and Short-Range Network using AI to Enhance Internet of Things," *J. Robot. Control (JRC)*, vol. 3, no. 2, pp. 505-518, 2022.

[33] G. P. N. Hakim, D. Septiyana, and I. Suwarno, "Survey Paper Artificial and Computational Intelligence in the Internet of Things and Wireless Sensor Network," *J. Robot. Control (JRC)*, vol. 3, no. 4, pp. 439-454, 2022.

[34] A. A. Sahrab and H. M. Marhoon, "Design and Fabrication of a Low-Cost System for Smart Home Applications," *J. Robot. Control (JRC)*, vol. 3, no. 4, pp 409-414, 2022.

[35] M. H. Widianto, A. Sinaga, and M. A. Ginting, "A Systematic Review of LPWAN and Short-Range Network using AI to Enhance Internet of Things," *J. Robot. Control (JRC)*, vol. 3, no. 4, pp. 505-518, 2022.

[36] N. S. Irjanto and N. Surantha, "Home Security System with Face Recognition based on Convolutional Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 408-412, 2020.

[37] A. Verma and M. K. Srivastava, "Real-time Face mask Detection Using Deep Learning and MobileNet V2," *VLSI, Microwave and Wireless Technologies, Lecture Notes in Electrical Engineering*, vol. 877, pp. 297-305, 2022.

[38] M. Sandler, A. Howard, M. Zhu, Andrey Zhmoginov, and L. C. Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *The IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, pp. 4510-4520, 2018.

[39] V. Patel and D.Patel, "Face Mask Recognition Using MobileNetV2," *Int. J. Sci. Res. Comp. Sci. Eng. Inf. Technol*, vol. 7, no. 5, pp. 35-42, 2021.

[40] H. Sun, S. Zhang, R. Ren, and L. Su "Maturity Classification of "Hupingzao" Jujubes with an Imbalanced Dataset Based on Improved MobileNet V2," *Agriculture*, vol. 12, no. 9, pp. 1-17, 2022.

[41] G. Edel and V. Kapustin, "Exploring of the MobileNet V1 and MobileNet V2 models on NVIDIA Jetson Nano microcomputer," *J. Phys. Conf. Series*, vol. 2291, no. 1, pp.1-7, 2022.

[42] C. L. Amar et al., "Enhanced Real-time Detection of Face Mask with Alarm System using MobileNetV2," *Inter. J. Sci. Eng.*, vol.7, no. 7, pp. 1-12, 2022.

[43] M. Sandler, A. Howard, M. Zhu, Andrey Zhmoginov, and L. C. Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *The IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, pp. 4510-4520, 2018.

[44] Y. Nan, J. Ju, Q. Hua, H. Zhang, and B. Wang, "A-MobileNet: An approach of facial expression recognition," *Alexandria Eng. J.*, vol. 61, no. 6, pp. 4435-4444, 2021.

[45] A. Srivastava1, A. Dalvi, C. Britto, H. Rai, and K. Shelke, "Explicit Content Detection using Faster R-CNN and SSD MobileNet v2," *Int. Res. J. Eng. Technol. (IRJET)*, vol. 7, no. 3, pp. 5572-5577, 2020.

[46] J. Jeong, H. Park, and N. Kwak, "Enhancement of SSD by concatenating feature maps for object detection," *arXiv preprint arXiv:1705.09587*, 2017.

[47] R. Redondo and J. Gilbert, "Extended Labeled Faces in-the-Wild (ELFW): Augmenting Classes for Face Segmentation," *Comput. Sci.*, 2020.

[48] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-shot Multi-level Face Localisation in the Wild," *2020 IEEE/CVF Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 5203-5212, 2020.

[49] A. Wirdiani, P. Hridayami, and A. Widiari, "Face Identification Based on K-Nearest Neighbor," *Sci. J. Inform.*, vol. 6, no. 2, pp. 150-159, 2019.

[50] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook," *IEEE Access*, vol. 8, pp. 220121-220139, 2020.

[51] P. Vrushali, and M. SanJay, "Autonomous Vehicle using Computer Vision and LiDAR," *I-manager's J. Embedded Syst.*, vol. 9, pp. 7-14, 2021.

[52] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj and L. Song, "SphereFace: Deep Hypersphere Embedding for Face Recognition," *2017 IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 6738-6746, 2017.

[53] S. Qureshi et. al. "Face Recognition (Image Processing) based Door Lock using OpenCV, Python and Arduino," *Int. J. Res. Appl. Sci. & Eng. Technol.*, vol. 8, no. 6, pp. 1208-1214, 2020

[54] N. Saxena, and D. Varshney, "Smart Home Security Solutions using Facial Authentication and Speaker Recognition through Artificial Neural Networks," *Int. J. Cognitive Comput. Eng.*, vol. 2, no. 2, pp. 154-164, 2021.

[55] M. H. Khairuddin, S. Shahbudin, and M. Kasssim, "A smart building security system with intelligent face detection and recognition," *IOP Conf. Series: Mater. Sci. Eng.*, no. 1176, 2021.

[56] M. R. Dhobale, R. Y. Biradar, R. R. Pawar, and S. A. Awatade, "Smart Home Security System using Iot, Face Recognition and Raspberry Pi," *Int. J. Comput. Appl.*, vol. 176, no. 13, pp. 45-47, 2020.

[57] T. Bagchi et. al., "Intelligent security system based on face recognition and IoT," *Materials Today: Proc.*, vol. 62, pp. 2133-2137, 2022.

[58] P. Kaur, H. Kumar, and S. Kaushal "Effective state and learning environment based analysis of students' performance in online assessment," *Int. J. Cognitive Comput. Eng.*, vol. 2, pp. 12-20, 2021.

[59] N. A. Othman, and I. Aydin, "A face recognition method in the Internet of Things for security applications in smart homes and cities," *2018 6th Int. Istanbul Smart Grids and Cities Congr. Fair (ICSG)*, pp. 20-24, 2018.

[60] T. V. Dang, "Design the encoder inspection system based on the voltage 7 wave form using finite impulse response (FIR) filter," *Int. J. Modern Phys. B*, vol. 34, p. 2040146, 2020.

[61] T. V. Dang and N. T. Bui, "Design the abnormal object detection system using template matching and subtract background algorithm," MMMS 2022, LNME, 2022.

[62] T. V. Dang and N. T. Bui, "Research and Design Obstacle Avoidance Strategy for Indoor Autonomous Mobile Robot using Monocular Camera," *J. Adv. Transp.*, 2022.

[63] T. V. Dang and N. T. Bui, "Multi-scale Fully Convolutional Network-based Semantic Segmentation for Mobile Robot Navigation," *Electron.*, vol. 12, no. 3, p. 533, 2022.

[64] T. V. Dang and Dinh-Son Nguyen, "Optimal Navigation based on Improved A* Algorithm for Mobile Robot," *Int. Conf. Intell. Syst. Netw.*, 2023.