IAAS: IoT-Based Automatic Attendance System with Photo Face Recognition in Smart Campus

Jaehoon (Paul) Jeong*, Minho Kim†, Yeonghyeon Lee†, and Patrick Lingga‡
* Department of Computer Science & Engineering, Sungkyunkwan University, Suwon, Republic of Korea
† Department of System Management Engineering & Consumer Science, Sungkyunkwan University, Suwon, Republic of Korea
‡ Department of Electrical & Computer Engineering, Sungkyunkwan University, Suwon, Republic of Korea
Email: {pauljeong, staraxer, ske03005, patricklink}@skku.edu

***Abstract*—The existing attendance checking system uses student identification (ID) cards or mobile applications. If the attendance is processed by an automatic attendance checking system, the attendance checking process can be more convenient for both students and professors than manual attendance checking. For getting these advantages, this paper proposes an IoT-based Automatic Attendance System (IAAS). IAAS is an attendance checking system using a face recognition technology. The image data of students are collected by a capturing device (e.g., smartphone and tablet PC) and processed by the face recognition system. This system checks who are attending the class and send their attendance to an attendance database system through an email notification.**

***Index Terms*—Internet of Things, Automatic Attendance Checking, Attendance System, Photo Face Recognition, Smart Campus**

I. Introduction

Nowadays, Internet-of-Things (IoT) technologies have been used in many different fields (e.g., smart home and smart campus) for providing convenience and efficiency to users. Similarly, this paper proposes an automatic attendance system different from existing attendance systems. The existing attendance systems are usually operated by a manual method or electrical system by identifying a student identification (ID) card or using an attendance application. There are some problems in these systems. For example, there is lecture time loss because of checking attendance by a manual method. Also, if students use an electrical attendance system, they always need to carry their student ID card or smartphone with an attendance application. A proxy attendance (i.e., attendance of somebody else for the sake of a person) can be done by an unconscionable person. This weak point of the existing attendance systems should be prevented.

This paper proposes an IoT-based Automatic Attendance System (IAAS) to facilitate an automatic attendance checking service by adopting a face recognition technology. This system can decrease any troubles generated in existing attendance systems. Because attendance is proceeded by user face data, the user does not need to carry any attendance equipment like an ID card or smartphone with an attendance application. Also, neither professor nor teacher does not need to check attendance manually because the attending process is progressed automatically. In addition, the IAAS includes an attendance reporting function such as an email transmission. Thus, the IAAS will be able to reduce attendance checking efforts and give convenience to both students and professors.

The main contribution of this paper is an improved automatic attendance system by analyzing several face recognition technologies. Through the performance evaluation of the IAAS and baselines, the benefits of our IAAS are shown.

The remainder of this paper is structured as follows. Section II summarizes the related work of face recognition. Section III formulates the problem of an automatic attendance checking service along with assumptions, and proposes a system architecture for such a service. Section IV describes the design of the IAAS using a photo face recognition technology. In Section V, the performance evaluation of the IAAS is presented against baselines. In Section VI, this paper is concluded along with future work.

II. Related Work

Our work is creating an automatic attendance system by utilizing Face Recognition technology. We use an existing face recognition technique, but enhanced them by comparing the existing technique and combine them to find the best result for our system.

A. SRGAN

Super-Resolution Generative Adversarial Network (SRGAN) is an algorithm that increases the image resolution [1]. SRGAN can produce photorealistic images. It consists of a Generative network and a discriminator network and can make a trained model for high-resolution photos for the given low-resolution photos as a train set. When a photo is given as an input through the trained model, a high-resolution photo is generated as an output. Another paper by Wang et al. [2] proposes Enhanced SRGAN (ESRGAN) to improve the performance by improving the network architecture, adversarial loss, and perceptual loss.

B. MTCNN

An input photo can be used to perform face detection through the MTCNN technique [3]. The MTCNN recognizes the face by varying the size of the image. The smaller the size of the image is, the easier it is for the box to find out contour features, and also the stronger the features become. When it is composed of P-net, R-net, and O-net, the input size increases in order. Initially, after reducing the scale, the face is detected through P-net. After that, the face detected on the P-net is scaled up on the R-net, and then a more sophisticated part is detected. On the last O-net, the scale is further increased for a feature [4]. In this way, face detection can be achieved in photos.

C. GoogleNet

A GoogleNet technique is used to make a Face Verification model [5]. Szegedy et al. propose a GoogleNet, which is a deep convolution neural network architecture named Inception. The main advantage of this method is a significant quality gain at a modest increase in computational requirements when its neural network is compared to shallower and less wide networks. In addition, the data are trained using the vggface2 data set. However, vggface2 provides a huge amount of data, but because it is a picture of Westerners, it may have the disadvantage of poor learning efficiency in Asians [6]. To solve this, this system can train a user's face data by transfer learning techniques.

III. Problem Formulation

This section presents our goal and assumptions and proposes an IoT-based Automatic Attendance System called IAAS. Given the related face recognition technique, our goal is to provide an automatic attendance system in university lectures by adding a face recognition technology in order to increase convenience for those who use the attendance system and prevent an illegal attendance activity, that is, proxy attendance. In other words, this paper aims at providing an improved automatic attendance system including suitable face recognition techniques.

A. Assumptions

Our objective is to provide an automatic attendance checking system using an instructor's smartphone rather than using an IoT surveillance/monitoring system. This work is based on the following set of assumptions on attendance environment settings. Multiple photos are required to identify who a user is. One of core elements of an attendance system is operational reliability. An attendance system should be able to recognize each person's face and identify who are attending now with a high accuracy. This process will be evaluated by the amount of system errors. Thus, this paper is expected that multiple photos of each user are available for the training of a proper automatic attendance system. Still photos with similar postures are given. To increase an identification rate, the IAAS should get multiple photos of students. Also, the face of each student is included with similar postures when the students use the attendance system.

B. Android Application and Firebase Storage

For face recognition, the IAAS takes group photos at a predetermined place. Though there are various capturing devices, we choose a smartphone with a camera and an attendance application. We make a capturing application suitable for our IAAS system with open source. There are two main functions of the application. The first function is to take a picture or select an image in a basic gallery application. After that, the second function is to send the collected image to our Firebase storage. With this procedure, we can test our system entirely. For example, the application collects group images by a smartphone camera and sends it to the Firebase storage. Our face recognition system gets the image in the Firebase storage and perform face recognition with that image.

C. IAAS Architecture

Our IAAS system's architecture is composed of three Parts such as Face Detection, Face Verification, and Notification. With a received image data, the IAAS system extracts each user's face data by MTCNN that is a face detection technology. With these extracted data, the IAAS system verifies each user to tell who the user is by GoogleNet that is a face verification technology. After the verification, the IAAS system notifies each user of the attendance completion by an email message.

IV. Design of IAAS Using Photo Face Recognition

A. Face Detection with MTCNN

With received images, the user's face data is extracted for face recognition. MTCNN is one of the face detection technologies that have strength in a wild environment. The proposed system constructs a face detection function with the MTCNN technology. An additional important thing is whether the ESRGAN technique is included or not. Thus, we tested the performance of MTCNN with ESRGAN and that of only MTCNN (i.e., MTCNN without ESRGAN), using the same test data.

ESRGAN is a technique to increase the image resolution as mentioned in [1]. While MTCNN is a method to recognize face that leverages deep convolutional networks and landmark location. The improvement of image resolution hardly improves the face detection of MTCNN, since the detection is based on landmark location of the face. Thus, the usage of ESRGAN technique can be ignored.

B. Face Verification

To select a suitable face verification model, we choose candidate models of face verification such as GoogleNet and VGG16. Like the above case, we tested each model with the same test data, and it is found that GoogleNet has better performance than others. We also checked which training method is suitable for the IAAS system. IAAS system trains face verification models with the vggface2 dataset. However, vggface2 is composed of Westerners, so it has low relation with the assigned users who are Asians. Thus, we selected a transfer learning method for the training model. In that case, we should prepare a suitable user's face data. To get the user's face data, we tested two cases. Each case proceeds with the received user images in advance. The first case is to extract a user's face data with the MTCNN technique and train it. On

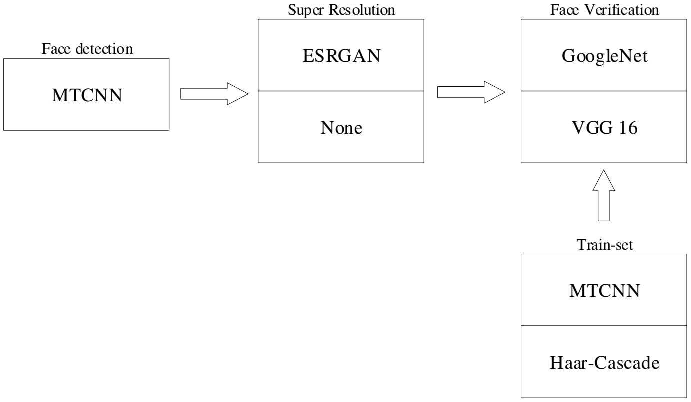

Fig. 1. Testing Structure

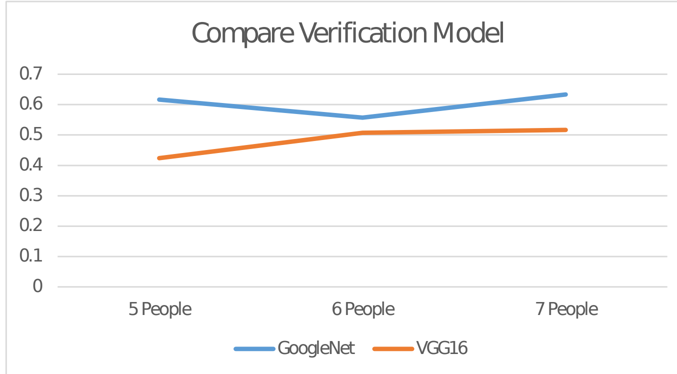

Fig. 2. Verification Models Comparison

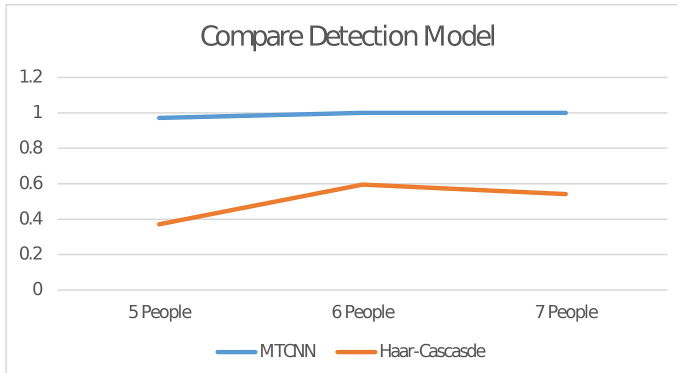

Fig. 3. Detection Models Comparison

the other hand, the second case is to extract the user's face data with the Haar-cascade technique and train it.

C. Notification Service

One of the important things of an attendance system is reliability. So, this system provides a notification service to the users. A notification service generates an attendance completion message when a user's attendance is detected by the the attendance system. To construct such a notification service, we use an email protocol that is Simple Mail Transfer Protocol (SMTP) [7]. Message Queuing Telemetry Transport (MQTT) for IoT message exchange can also be used for this notification [8]. The message delivery using MQTT is left as future work.

We construct a database (DB) with Firebase for storing user information. The DB is constructed with Lecture, Username, User email, and so on. With this pre-constructed DB, whenever it verifies each user, the system then checks the attending user's data and finds the user's email address. Lastly, the system sends the attendance completion message to the attending user. In addition, to increase reliability, the notification message is composed of a user's name and time information.

V. Performance Evaluation

A. Construct Testing cases

Fig. 1 is the testing structure of the IAAS system. To find a best composition for our IAAS system, we construct testing cases with related processes. First, we divide super-resolution cases according to the inclusion of ESRGAN. Second, we construct two face verification models such as GoogleNet and VGG 16. Lastly, we divide the getting train-set method such as MTCNN and Haar-Cascade.

With these testing systems, we evaluate each case by the same testing data. The testing of data goes through each case and returns a face recognition message to an attendant who is attending now. With that message, we evaluate each case how the amount of each case can recognize each user accurately. In other words, we evaluate each case how amount the system can verify each user rightly with the same testing data.

B. Testing Environment

For proof of concept (POC), we used a small size of samples. Even with this small sample size, we showed the effectiveness of our scheme. We constructed the proposed testing environments. Before testing, we chose candidate users and prepared group photos. Also, we collected the selected users' data and train them with the proposed two methods such as MTCNN and Haar-cascade. After that, with the grouped photos, we tested each case and obtained results.

C. Testing Result

We tested face verification models by two cases (i.e., GoogleNet and VGG 16) as shown in Fig. 2. We selected five candidate persons. The collected group pictures have 5, 6, and 7 persons. With these group pictures, we tested each model with the same cases. With this test, it is found that GoogleNet showed better performance than another. The result shows that the average verification accuracy of GoogleNet was about 0.595 and the confidential interval was about (0.521 $\sim$ 0.670) and the average verification accuracy of VGG16 was about 0.475 and the confidential interval was about (0.371 $\sim$ 0.578). Thus, we selected the case without ESRGAN and with GoogleNet for face verification.

Second, we tested two cases (cropped by either MTCNN or Haar-cascade) for selecting the training method. As can be seen in Fig. 3, it is found that Haar-cascade showed better performance than MTCNN. The result shows that the average detection accuracy of MTCNN was about 0.988 and the confidential interval was about ($0.966 \sim 1$) and the average detection accuracy of Haar-Cascade was about 0.494 and the confidential interval was about ($0.366 \sim 0.622$). It is concluded that this result is caused by different crop methods for each model. The image sizes cropped by MTCNN are different from each other. However, the image sizes cropped by Haar-cascade are fixed.

VI. Conclusion

This paper proposes an IoT-based Automatic Attendance System (called IAAS). Our goal is to provide a reliable and improved automatic attendance checking system with face recognition technologies. This goal is achieved by constructing an improved system structure with suitable face recognition technologies and using training data with Haar-cascade to extract a user's face data. With this system, it is expected that that our IAAS system can provide convenience for attendance checking to both students and professors. As future work, we will explore the impact of postures or appearance changes of users on the performance and develop a solution to deal with such factors.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2017-0-01633) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). This work was supported in part by IITP grant funded by the MSIT (No. 2019-0-01343, Regional strategic Industry convergence security core talent training business). Note that Jaehoon (Paul) Jeong is the corresponding author.

References

[1] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105–114.

[2] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, September 2018.

[3] Y. Fu, M. Kim, and J. wook Jang, "A Study on Face Detection Algorithm Implementation Based on MTCNN Model for Complex Environments," in *Korea Institute of Information and Communication Engineering (KIICE) Fall Conference*, vol. 23, October 2019. [Online]. Available: http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09262354

[4] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[5] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[6] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[7] J. C. Klensin, "Simple Mail Transfer Protocol," *IETF RFC 5321*, October 2008.

[8] R. Toasa, C. A. S. Villafuerte, D. Martinez, J. L. V. Zambrano, and C. Gallardo, "Secure Push Notification Service based on MQTT Protocol for Mobile Platforms," in *JIISIC 2017*, August 2017.