J Inf Process Syst, Vol.15, No.4, pp.904~919, August 2019
https://doi.org/10.3745/JIPS.03.0127

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

CCTV-Based Multi-Factor Authentication System

Byoung-Wook Kwon*, Pradip Kumar Sharma*, and Jong-Hyuk Park*

Abstract

Many security systems rely solely on solutions based on Artificial Intelligence, which are weak in nature. These security solutions can be easily manipulated by malicious users who can gain unlawful access. Some security systems suggest using fingerprint-based solutions, but they can be easily deceived by copying fingerprints with clay. Image-based security is undoubtedly easy to manipulate, but it is also a solution that does not require any special training on the part of the user. In this paper, we propose a multi-factor security framework that operates in a three-step process to authenticate the user. The motivation of the research lies in utilizing commonly available and inexpensive devices such as onsite CCTV cameras and smartphone camera and providing fully secure user authentication. We have used technologies such as Argon2 for hashing image features and physically unclonable identification for secure device-server communication. We also discuss the methodological workflow of the proposed multi-factor authentication framework. In addition, we present the service scenario of the proposed model. Finally, we analyze qualitatively the proposed model and compare it with state-of-the-art methods to evaluate the usability of the model in real-world applications.

Keywords

Argon2, Convolutional Neural Network, Deep Reinforcement Learning, Physically Unclonable Functions

1. Introduction

CCTV (closed circuit television) is one of the most valuable devices connected to IoT applications such as transportation, healthcare, education and so on. The connection of these devices encourages us to transform our lifestyle today by optimizing efficiency and improving customer service [1]. Most video surveillance systems use cameras to convert the data of objects into image data, and then store them in the video recorder known as DVR (digital video recorder). Many users can access the video recorder through a web browser or the Internet on a remote system. The use of DVR in IoT applications involves accessing the organization's network; thus increasing the risk in the CCTV system exponentially. Note, however, that many users do not recognize these risk or dangers in IoT applications. IP cameras and live video search engines can be infected with malicious code, and people with malicious intent can easily use it [2,3]. Therefore, if the CCTV system does not have any additional security features, it will pave the way for launching various attacks in many applications.

According to previous studies, the latest CCTV systems based on DVR and cloud systems are vulnerable to a wide range of cyber-attacks, which can serve as a potential starting point for damage or

Manuscript received April 9, 2019; first revision July 29, 2019; accepted July 31, 2019.
Corresponding Author: Jong Hyuk Park (jhpark1@seoultech.ac.kr)
* Dept. of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech), Korea ({rnjsqud123, pradip, jhpark1}@seoult.ac.kr)

extraction of network information. That is because DVR-based CCTV systems use ports to provide general access. This allows browsers to connect efficiently through firewalls. Many DVRs automatically configure ports to be forwarded and automatically ask the router to forward the port without notifying the user. The DVR-based CCTV system's configuration is easy but not fully secure because it has some limitations. To address these limitations, CCTV uses the digital video surveillance system, which is capable of capturing images and video that can be compressed and stored in the IoT communication network. Currently, the CCTV system is a video surveillance system, and it is continually spreading in various fields. In addition, it monitors the surrounding situation with a simple function. Nowadays, it is growing rapidly as a network-based intelligent CCTV [4]. The CCTV system can automatically detect and track all the characteristics of objects or persons among the images captured by the CCTV camera, so the surrounding situation monitoring function is used very effectively.

Among them, video security is being used to build an accurate traffic control system through the identification of traffic violations and traffic accidents. A crime prevention system is established through the identification of appearance and criminal behavior of criminals, whereas a disaster prevention system is set up through the identification of fire and terrorist act. There are advantages and disadvantages in the field of video security. Advantages include the enhancement of performance and quality of video security equipment, improvement of technology awareness and interest, expansion of coverage and utilization of CCTV-dedicated hardware and software, development of various functions and improvement of performance, and development of related fields. Rapid increase in manpower and spread of technology are taking place. The disadvantages are the added complexity compared with the existing analog method and need for retraining. It is also necessary to improve the technical security of hardware and software and equipment according to the interworking of various devices.

The user authentication system is based on the identity of the user who requested access to the resource in the real application as a process of accessing a particular system but verifying the identity of all users whose access is controlled as the mechanism. File permission, data permission, and program permission are part of user authentication and access control. Initially, user authentication is performed based on the use of passwords or tokens but has fundamental limitations such as security and usability. In other words, a short-length cipher or a password generally has low entropy, so an attacker can guess it easily. On the other hand, since a long cipher has high entropy, it is hard to remember for the attacker. It is also difficult for users to remember long and secure passwords for each employed service. This causes the same or similar passwords to be used for each service, greatly increasing the risk of compromising passwords and corrupting related services. On the other hand, tokens can be easily misplaced or stolen. Next-generation authentication technology such as Blockchain [5] and biometrics [6] is rapidly growing in authentication fields. With the development of next-generation authentication technology, authentication methods and technological innovations are progressing by applying various platforms such as mobile and cloud. In addition, innovative technologies and services are required to cope with market changes due to the spread of private authentication service markets such as messenger-based authentication service (SNS) and simple financial authentication service.

To address this security problem, two-factor authentication (2FA) was used in places where IoT devices were used as the second authentication factor [7]. The process involves authenticating a user using two authentication factors and granting access to the system [8]. It expands the mechanism by adding a step to the 2FA-based authentication process [9] that users can access through dynamically generated disposable tokens. It provides 2FA that satisfies both single-factor authentication and high security [10]. Among other methods, biometric data such as fingerprints or retina are used as the second authentication

factor [11].

In 2016, China Central Television reported the communication fraud for dual certification. Criminals using dual authentication could easily steal victims' assets through China Mobile's online 4G card replacement [12]. 2FA can identify the user's identity by using a combination of three elements, such as smart card [13], mobile phone [14], and fingerprint. For example, the dynamic token method provides high security with a one-time encryption scheme. When visiting other sites, however, the user must fetch another token.

With the development of various kinds of wireless infrastructure technology, network communication connection can be used easily without time and space restrictions. As such, there is a tendency for the range of network communication to be narrowed based not only on specific time and space but also on a wireless network, such as connection between people and objects and between objects. Due to the influence of the Internet environment, small-sized objects are integrated with computing technology to form a network for all connectable elements, and various technologies are fused to form a specific service. There are various security threats due to the weakness of technology itself or implementation method. Especially, Internet services that send and receive various protocols and massive amounts of data are exposed to security threats, medium attacks, data strikes, tampering, spoofing, and eavesdropping. These security threats require efficient, enhanced dual authentication services to prevent unsafe Internet service environments.

The main characteristics of this paper are as follows:

- We propose the three-step user authentication multi-factor security framework, which uses commonly available and inexpensive devices such as onsite CCTV cameras and smartphone camera to provide fully secure user authentication.
- The proposed framework provides superior detection accuracy for user image recognition using deep learning reinforcement-based convolutional neural network (CNN).
- To ensure that the server receives image data from an authorized smartphone device only, the proposed model has used physically unclonable functions identification (PUF-ID). It prevents an attacker from injecting false data using unauthorized devices.
- We also propose hashing the features extracted from both primary and secondary images using a superior Argon2 hashing algorithm. We compare the hash values of both images to ensure complete security in the event the CNN fails to identify a malicious authorization attempt.
- We discuss the service scenario of the proposed model. Lastly, we analyze qualitatively the proposed model and compare it with state-of-the-art methods to evaluate the usability of the model in real-world applications.

The rest of this paper is organized as follows. Section 2 discusses the backgrounds of preliminary and related research. Section 3 proposes a three-phase user authentication multi-element security framework and discusses the methodological workflows. Section 4 analyzes the advantages of the proposed 3-step authentication framework. Section 5 presents the conclusion of the study.

2. Background

2.1 Preliminaries

Authentication is a process of validating credentials such as user password, user id, and OTP (one-time password) to verify identity in any application [15]. Among the existing authentication technologies are

methods such as login-password authentication and OTP through object information. Technically, object-based login authentication is the simplest, but the password itself can be easily exposed. An object usually decides a password by using keywords and symbols related to it, and an attacker can easily steal a password through a preliminary investigation of the attack target and a social engineering technique. The unique ID of a device on the Internet must be used in conjunction with other authentication methods because it is easy to duplicate. OTP, which is known as the most powerful authentication method, should be automatically authenticated between devices due to the characteristics of the object Internet environment.

The identity of an individual is determined by one thing and another thing, and at least two or three authentication elements related to security should be verified in the system. The authentication factor may differ from one of the following depending on the level of security:

- One-factor authentication (1FA): The simplest authentication method for granting user access to a system, such as a network or a website, is password.
- Two-factor authentication (2FA): Information that the user knows (name, password, etc.) is used in a two-step authentication process.
- Multi-factor authentication (MFA): This is an independent and the most advanced authentication concept for granting users access to more than one element in systems requiring security.

Implementing a MFA system increases security by requiring the user to provide an additional set of credentials before being granted access because single-factor authentication system (username and password) is no longer safe enough, and neither are 2FA systems fully secure. As a device with the existing dual authentication technology, an ATM device can perform transactions through a proper combination of bank card and PIN. If the credentials are stolen, a user can be impersonated. MFA is a more secure concept compared to 1FA and 2FA. Used in many applications such as smart city, healthcare, transportation, smart farming, and so on, MFA can also be used to control access to AWS service APIs.

Deep reinforcement learning is a method wherein an autonomous agent accelerates neural network design using trial-and-error algorithm and cumulative compensation function. Deep learning (supervision, class supervision, non-supervision) is part of machine learning and is based on the layers used in artificial neural networks. As the thing in the Internet environment, reinforcement learning serves as the foundation of automation edge applications and contributes in a variety of environments. By using the trial-and-error approach, reinforcement learning helps decision makers draw up action to help the machine and achieve an optimal algorithmic model for random conditions. It accepts and acts as a systematic way of continually rewarding or penalizing the results to improve machine decision making.

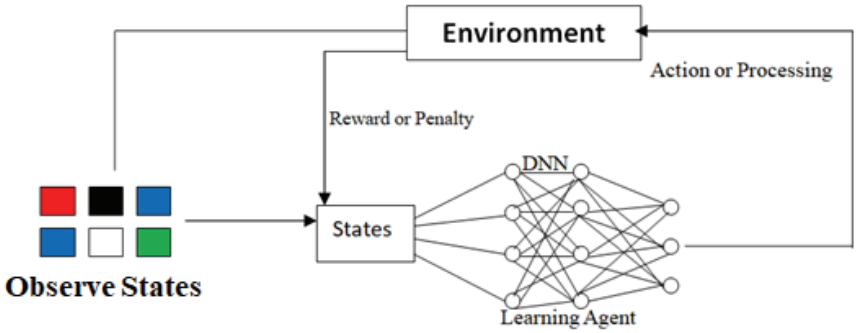

Fig. 1. Deep reinforcement learning.

Deep learning reinforcement learning has been used for face recognition mechanisms under the CNN's facial feature analysis for facial feature extraction and is converted into another comparison for user identification. Fig. 1 presents the generic workflow of deep reinforcement learning. Wang et al. [16] used the feedforward network architecture for the CNN and proposed an authentication technique that uses a backpropagation algorithm to improve the recognition rate of face recognition effectively.

PUF-ID stands for physical unclonable function identification; it is like a physical defined digital fingerprint that serves as a unique identification for integrated chip-related devices like a microprocessor. Used for high security and privacy requirements, specifically cryptography, this technology enables extracting security keys and unique identifiers according to the inherited characteristics of each integrated chip. PUF utilizes the submicron deformation that occurs during chip production, giving each transistor a little electrical character and a unique identity. Two types are used: strong PUF and weak PUF. A strong PUF generates a large number of challenge-responses, whereas a weak PUF generates a limited number of challenge-responses. Practically, powerful PUF is used because it enhances security in applications. An Brussel. [17] have proposed it for efficient non-negotiation in the communication phase between IoT devices connected to the same authentication server, and it provides authentication and non-repudiation when using elliptic curve multiplication.

As a cryptographic hashing algorithm, Argon2 is used for securing the password in any application. It is mainly designed for configuration runtime, and memory consumption means how long it takes to hash the password and how much memory is required for the user. Argon2 has three subcategories: Argon2d, Argon2id, and argon2id. Argon2d is better suited for password-caching applications used to access data-dependent memory with faster algorithms and is less susceptible to side-channel attacks, but is not suited for cryptographic hashing. Argon2i is used for independent data memory access and cryptographic hashing using cryptographic keys. Argon2i is slow because it consumes a lot of memory to protect data from tradeoff attacks. Argon2id is a cryptographic hashing algorithm that differs from the existing Argon2i and is a combination of independent property-dependent properties on data. Argon2i's resistance to cache channel attacks and Argon2d make the GPU resistant to cracking attacks. Sharma and Park [18] proposed a structure for distributed and smart city networks based on the Argon2 proof-of-work (PoW) scheme in smart city IoT applications. To solve some problems such as raw recalculation attacks, the MTP-Argon2 hash algorithm is used to improve memory strength by storing parallel search and combinable attack patterns in the memory [19]. Biometrics is not a complete authentication method but is very effective as one of the authentication technologies that identify objects by providing secure, convenient authentication. The biometrics capabilities of these devices provide a secure ecosystem that uses biometrics data unobtrusively utilizing integrated hardware channels, allowing users to use most personal devices such as cell phones conveniently and securely. This method involves capturing the user's fingerprints, extracting functions, comparing information based on the database, and outputting the authentication results. Nowadays, biometrics is recognized as a safe method among mobile device authentication methods. Biometric information that can be used for such authentication is owned by the individual and must be personally identifiable and non-reproducible. Although biometric technology can provide more security than traditional password-based security systems, there are certain limitations due to some technical barriers. For example, identification sensor devices with unstable recognition rate with regard to biometric information and high recognition rate are generally expensive, so it will take time before it can be used more easily in various fields. Since biometric information is personal data that causes serious damage in case of leak, systematic management system and relatively high cost are required.

Therefore, in order to use biometric information as a security measure, a management system for preventing sensitive personal information infringement must be considered. Other types of information such as geographic location and time may be further included in the authentication process. Note, however, that at least two of the three factors above should always be used. Location information and time data may be used to limit remote access to the entity network according to the individual work schedule.

Image security can be defined as an area of physical security that acquires image information for the purpose of monitoring using a video camera such as a camera or an image storage device and transmits the image information to a specific user. Examples of the components include an input device such as a camera, a display device such as a monitor, a video storage device such as a DVR, and other peripheral devices. The scope of the video security industry includes companies involved in the production, installation, and delivery of security equipment related to CCTV, and related fields include security/ drone, CCTV, related industries (black box, smart home, IoT equipment, etc.). The video security technology can be classified into two categories according to purpose: the background pretreatment technique including background modeling and image correction, detection, and classification technology for detecting and classifying the objects of interest and tracking technology for continuously tracking the detected objects, and; security technology to prevent security dysfunctions.

2.2 Related Works

Many researchers have studied user authentication using many technologies. Wang et al. [20] proposed an edge-based face verification system for privacy protection. Edge computing is used in the authentication system for privacy verification, with the CNN utilized in facial feature extraction. Lin et al. [21] proposed a face recognition system with user authentication using a support vector model classifier. Based on facial features, the online face recognition system provides enhanced local binary pattern in spatial part by training the SVM classifier and solves the problem of classification accuracy. The facial attributes proposed by Samangouei et al. [22] for active authentication on mobile devices, wherein facial attributes are used for the authentication of smartphone users, train a bunch of binary attributes classifiers used for the visual descriptions of faces, and learned classifiers are applied to the image of the current user of the mobile device to extract the attributes; authentication is then completed by comparing the calculated attributes with the enrolled attributes of the original user. In the CNN-based anti-spoofing two-tier MFA system proposed by Sajjad et al. [23], tier1 integrates the hash of fingerprint, palm vein, and face recognition to match with the corresponding databases, and tier2 is used for fingerprint, palm vein print, and face anti-spoofing CNN-based model to detect spoofing. Lee et al. [24] proposed a remote anti-spoofing scheme for the face authentication system that provides a systematic approach to face detection incorporating state-of-the-art liveness detection and face verification algorithm to safeguard a system against attacks. Face liveness detection has feature extraction, classifier training, and face verification as subparts. Zhang et al. [25] proposed a conventional username and password-based single-factor authentication scheme, but it has a limitation, i.e., vulnerability to dictionary attacks, snooping, and brute-force attacks. Zhu et al. [26] proposed a learning-based scheme to authenticate the ownership of mobile devices in real time. They conducted a feasibility analysis of the proposed system on smartphones and smartwatches environment. A MFA key exchange protocol for mobile communications was presented by Zhang et al. [27]. Ometov et al. [28] cited the challenges of the

MFA scheme for securing IoT applications. They also discussed the advantages and disadvantages of various approaches.

3. Proposed Framework

In this section, we discuss the proposed multi-factor face authentication framework. To the best of our knowledge, we are the first to introduce a complete and secure framework for image authentication. The benefits of the proposed scheme are as follows: (1) it favors a rigorous 3-step process to authenticate a user, and (2) it offers a completely secure environment with no room for erroneous or inaccurate authentication check.

3.1 Overview of Multi-factor Authentication

Fig. 2 presents the proposed multi-factor face recognition that consists of four main steps: (1) face detection and recognition of primary image captured by CCTV; (2) upload, detection, and face recognition of the secondary image uploaded by the user using a secure smartphone; and hashing of both images, and (4) comparison of hash values of both images. Note, however, that the existing methods are not sufficient for a complete and secure environment as they rely entirely on AI-based solutions characterized by incorrect face authentication.

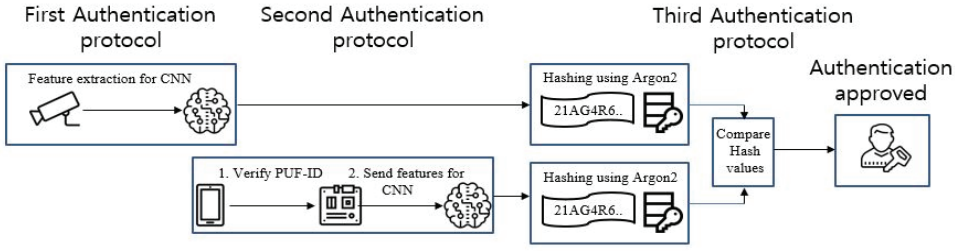

Fig. 2. Multi-factor authentication framework.

This study proposes the use of the following technologies to provide a secure framework: (1) face detection and recognition using deep reinforcement learning, which will study and analyze the selected features to provide the first authentication for both primary and secondary images; (2) PUF-ID is used to ensure that each device has its hardware-embedded digital fingerprint, no single device has the same identifier as the other, and PUF-ID cannot be altered through cryptographic means but requires a change at the physical level; and (3) Argon2 cryptography offers superior hashing technology with outstanding performance and security compared to other hashing methods and allows for the customization of duration to hash the data and the memory required to hash successfully.

3.2 Methodological Workflow

In our proposed framework, we have the security application placed on the cloud server. The deep reinforcement algorithm process occurs at the cloud as well as the hashing process. The dataset implemented will contain features that are consistent for any given user and which include details such

as the distance between the eyes, depth of eye sockets, length of the jaw line, width of the nose, and shape of the cheekbones. The algorithm will be trained using the aforesaid features. The user's smartphone has a PUF-ID that is unique to it and is stored at the server to be used as part of the verification method wherein data is received from the authorized device. The methodological workflow of the framework is shown in Fig. 3, and it operates in the following 4-step process:

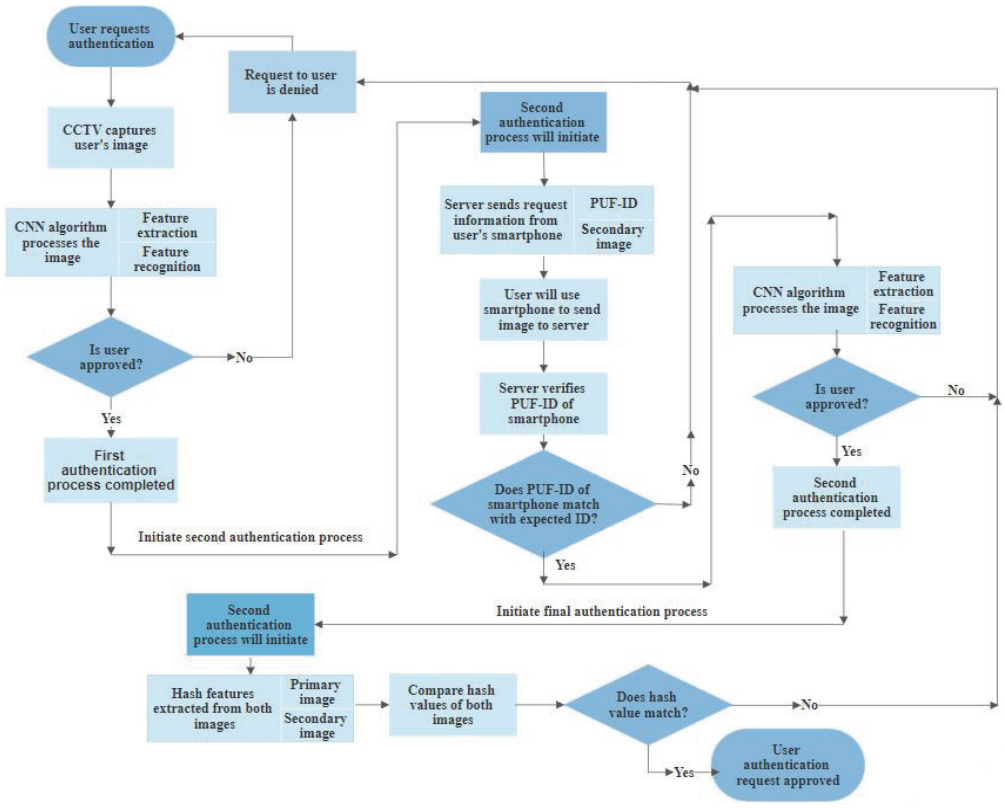

Fig. 3. Methodological workflow of the proposed framework.

- Step 1: CCTV will capture the user's photos upon request for authentication and will send them to the security application. The CNN algorithm is formed from the features present in the dataset. This step begins with extracting the features contained in the captured image by CCTV. The next step is to run the test set of the algorithm on the training set to determine whether the user is allowed to continue or not. Once the algorithm identifies and approves the user, the first step in the authentication process is complete.
- Step 2: The second authentication process begins with the server asking the user to take a photo using his/her smartphone. This smartphone has a unique digital fingerprint using PUF-ID technology. The device ID will be sent to the server with the image for secondary authentication. The server will then compare the ID with the expected device ID and, after verification, allow the extraction of the features of the image to execute the CNN algorithm. Once the authorized user is identified and recognized, the second authentication method is complete.
- Step 3: The server using the Argon2 hashing algorithm will hash the image features of the two images separately.

- Step 4: The final authentication protocol will result in a comparison of the two hashes in order to ensure that the features of the two images of the user match. The image can be processed under different conditions, e.g., if the user modifies his/her expressions; note, however, that his/her facial structure remains the same. Comparing the hash values will confirm that these two users are in fact identical. Finally, the user will be authenticated successfully, and access will be granted.

Algorithm 1 presents the MFA algorithm. The proposed MFA framework ensures the highest security possible for face recognition in an environment requiring high security. Instead of relying solely on AI-based algorithms for final security, our proposed model takes into account the possibility of second-factor authentication emerging from a malicious user's device. The use of PUF-ID technology ensures that only the authorized user will have access to send the secondary image for the second authentication process. Furthermore, the use of hashing of facial features instead of the images themselves ensures that any malicious user with similar facial features cannot gain unlawful authorization. Photo lighting conditions and user expressions may change, but the facial features remain constant; as such, the hash value of the authentic user will always match. The proposed framework presents a fully secure environment for the face authentication of users.

Algorithm 1. Multi-factor authentication

Input: CCTV_Image ← Capture User Image, Smartphone_Image ← Capture User Image, Smartphone_PUF-ID ← PUF-ID, Flag1 ← False, Flag2 ← False

Begin

Phase 1: CCTV
FeatureExtraction ← CNNAlgorithm (CCTV_Image)
FeatureRecognition ← CNNAlgorithm (FeatureExtraction)
If CCTV FeatureRecognition matches Stored Features **Then**
Set Flag1 ← True
Message "Authorization success! Proceed to take picture using smartphone camera"
Else
Message "Access denied. Try again"

Phase 2: Smartphone
FeatureExtraction ← CNNAlgorithm (CCTV_Image)
FeatureRecognition ← CNNAlgorithm (FeatureExtraction)
If Smartphone_PUF-ID matches Stored PUF-ID **Then**
If Smartphone FeatureRecognition matches Stored Features **Then**
Set Flag2 ← True
Message "Authorization being processed. Please wait"
Else
Message "Authorization denied."
Else
Message "Authorization denied."

Phase 3: Hash
PrimaryHash ← Hash (Primary Image Features)
SecondaryHash ← Hash (Secondary Image Features)
If PrimaryHash Equals SecondaryHash **Then**
Message "Access granted**"**
Else
Message "Authorization denied**"**

End

3.3 Service Scenario

The service scenario of the proposed framework model is shown in Fig. 4. The black line shows the flow of normal usage patterns. When a service request arrives, it is handled in accordance with the three module procedures in the security model, and the service provider learns the normal usage pattern information of the providing application in order to assist in the initial setup. The Security Response Center helps evaluate whether a new attack has occurred or the security module has responded appropriately. The security module records the correspondence and connects with the Security Response Center to send information about the new security attack and to receive feedback from the corresponding information.

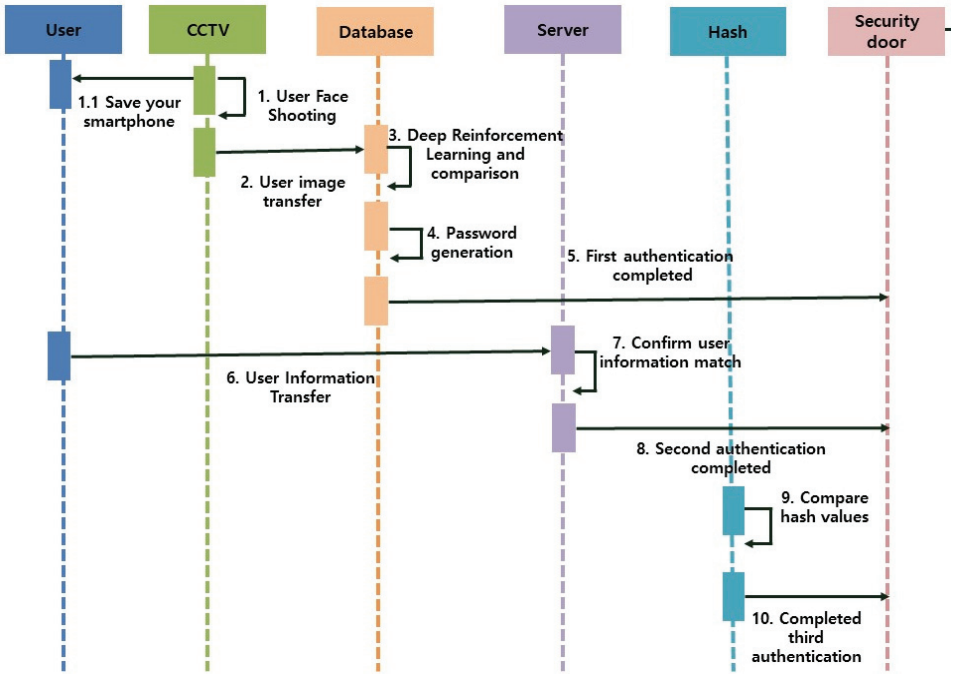

Fig. 4. Service scenario of the proposed two-factor authentication system.

In the proposed model, learning is performed on the initial data, and primary authentication, secondary authentication, and tertiary authentication occur. Initially, the first authentication is a method of authenticating the facial feature extraction and recognition through deep reinforcement learning after the face of the user is recognized. The facial features of the extracted user are learned to prevent intrusion by malicious users. Second, PUF-ID is used to check whether the hardware of each device has a built-in digital fingerprint. One device does not have the same identifier as another device. PUF-ID cannot be changed through encryption, so it must be changed at the physical level. Third, Argon2 encryption provides superior hashing technology with outstanding performance and security compared to other hashing methods. Argon2 allows customizing the duration to hash the data and memory needed for the hash.

Secondary authentication is a method wherein the secondary data is unlocked through the user's IoT device by hashing the initial data. The security countermeasure against the database storing the face data

utilizes the existing security threat database and improves the security service performance by updating the result and learning information after the correspondence. The analysis includes static and dynamic analysis, and the analysis result and learning are stored in the report. In the initial setting, sample data is collected to detect normal usage pattern, suspicion, and threatening behavior pattern through learning. The detection module ensures that service data classification can be done accurately and quickly through map learning. The defense module uses an in-depth learning algorithm to enable responding appropriately to the detected threats and to update log records and countermeasures against unknown attacks so that they can respond flexibly to the next attack.

4. System Analysis and Discussion

In this section, we discuss the proposed MFA framework in comparison with other state-of-the-art methods using facial image recognition as the primary source of user authentication. We have studied only the proposed methods from the past 3 years to exclude any obsolete research solution. A summary of the comparison study is presented in Table 1. The system analysis of the proposed framework is conducted based on three important criteria.

Table 1. Comparison of the proposed method with state-of-the-art methods

Study, year	Algorithm used	Technique used	Platform	Security	Operational complexity	Ease of use
Azimpourkivi et al. [25], 2017	Random forest and multilayer perceptron	Uses any smart device for user image identification and requires any user-preferred secondary device for secondary authentication	Smart devices/ Trinket	Low	Medium	Low
Sajjad et al. [23], 2018	Convolutional neural network	Compares the hash of fingerprint as primary authentication and palm vein and image as secondary authentication	Fingerprint scanner/ camera	Medium	High	Low
Wang et al. [20], 2019	Convolutional neural network	Cameras present in the building send user image to the edge layer for user authentication	Camera	Low	Low	High
Samangouei et al. [22], 2017	Support vector machine	Continuous face authentication-based system using the smartphone's front camera	Smartphone	Low	Low	High
Multi-factor authentication	Convolutional neural network	Images from CCTV and smartphone are authenticated using CNN. PUF-ID ensures that only the authorized device can send data. Hashing using Argon2 ensures complete security	CCTV/ smartphone	High	Low	High

- **Security:** The prime feature of the framework is to provide a completely secure environment for the system to detect accurately and avoid any spoofing attempts by malicious users. The proposed method uses images captured from CCTV and smartphone devices and utilizes the CNN to ensure the highest accuracy in detecting the authorized user. In the event a malicious user successfully manipulates the first authentication system, he/she will face the difficulty of producing a secondary image from an authorized device. A smartphone with a pre-registered PUF-ID at the server is allowed to send the secondary image for subsequent authentication, and no other device can communicate with the server. This step protects the server from any cyber-based attacks. A hashing algorithm using Argon2 is used to ensure that both the primary and secondary images sent have matching attributes or features that belong only to the user. Wang, et al and Samangouei et al. [22] both proposed a single authentication system using the CNN as the sole defense mechanism to detect any spoofing attacks by a malicious user. These proposed systems are not entirely secure and can easily be circumvented by a malicious user. As presented in Table 1, we ranked the models into low-, medium-, and high-security levels depending on the level of authentication factors and the possibility of compensating for the weakness of the model. Sajjad et al. [23] proposed the use of the Two-Factor security system using fingerprint-based security as the primary defense mechanism and palm vein and image as the secondary mechanism. They implemented a hashing method to compare the hashes of the fingerprint; note, however, that using clay or simple dental mold allows the attacker to bypass the fingerprint scanner easily. The palm vein can be easily replicated using the same method. The proposed method by Azimpourkivi et al. [25] implements a two-factor authentication system but uses an image-based approach only once. In the event a malicious user tricks the system, it is faced with another image recognition using a user-chosen device, which may be a watch, a doll, or any other item. Note, however, that it is easy for an attacker to be aware of the user's behavior through simple observation. Thus, the attacker can gain authorization without any difficulty.
- **Operational complexity:** Our proposed multi-factor authorization system requires no specialized training or tutorial for the user to operate. A CCTV camera is available onsite, and a user only needs his/her picture to be taken. A smartphone camera is used for the secondary image. The system proposed by Wang et al. [16] and Samangouei et al. [22] requires no specialized training as the user requires his/her image to be captured by the CCTV and a smartphone. The defense system proposed by Sajjad et al. [23] requires a high level of training for the user. Specifically, the user needs to learn the sequence of events first to scan his/her fingerprints, and then place his/her palm vein on a scanner device and finally have his/her image captured for the security system to function fully. The proposed system of Azimpourkivi et al. [25] requires basic training for the users to understand that they require a smart device and a trinket, which is an object of their choice. The need for the use of specialized hardware at all times must be instructed to the user.
- **Ease of use:** The proposed MFA system uses CCTV as the first point of collecting a user's image for security screening. This device is inexpensive to procure for any organization. Secondary authentication uses a smartphone, which is carried by nearly every user. No other hardware that the user may have difficulty with or require especially for security is needed. The use of a smartphone device ensures that the user will not forget to carry it with him/her, unlike the case of using any specialized hardware to gain authorization. Wang et al. [16] and Samangouei et al. [22]

proposed the use of a camera and a smartphone device, respectively. The camera can be webcam or CCTV, which are both inexpensive. On the other hand, all users generally carry a smartphone with them at all times. The defense system proposed by Sajjad et al. [23] requires separate hardware for detecting fingerprints and another for detecting palm veins. The proposed devices are expensive to procure compared to our proposed MFA system. The proposed system of Azimpourkivi et al. [25] requires a user-selected device that must be carried mandatorily at all times. In the event that a user forgets to bring the equipment, he/she will be unable to gain security authorization.

As discussed above, the principle of MFA is that each factor compensates for the weakness of the other factors. We can supplement it by adding an authentication factor that is difficult to guess. One of the main advantages of MFA is the significant reduction in the risk of end-user identities being compromised. The proposed secure multi-factor framework for user image verification via CCTV and Smartphone looks promising in terms of providing additional security. In addition, the operational complexity and ease of use features make the model applicable to real-world applications.

5. Conclusion

Nowadays, IT complexity is a real and ongoing issue. Every change to the network can potentially trigger a chain reaction of tweaks and adjustments that can inconvenience users and keep them offline. Because a simplified authentication process keeps the productivity (and morale) high, IT administrators need to ensure that every new upgrade or addition has as little effect as possible on access to critical programs. Using MFA, administrators can adapt the level of support needed to the contextual information, such as connection behavior patterns, facial reorganization, and type of connection system used. Note, however, that the principle of zero trust, i.e., users with a simple password for authentication should not be trusted without strong multifactor authentication, should be reinforced to ensure authorized human access.

In this study, we proposed a completely secure multi-factor framework for user image verification via CCTV and smartphone. Using deep reinforcement learning-based CNN algorithm, we ensured that the system offers the highest accuracy in detecting the accurate facial features of the user. PUF-ID technology is used to ensure that only authorized devices can send the image to the server for secondary image verification. The hashing of the features of both primary and secondary images and their subsequent comparison ensure that no malicious user can deceive the system and gain access. We also conducted a qualitative analysis of the system model to assess the superiority of the proposed model.

In the future, we will expand the proposed model to integrate it with the edge computing network to provide more resource-efficient, low-latency services to a scalable IoT network.

Acknowledgement

This study was supported by the Advanced Research Project funded by the SeoulTech(Seoul National University of Science and Technology).

References

[1] R. Dobson, "IoT security for connected surveillance cameras," 2016; https://www.iotsecurityfoundation.org/iot-security-for-connected-surveillance-cameras/.

[2] V. Suryani, S. Sulistyo, and W. Widyawan, "Two-phase security protection for the Internet of Things object," *Journal of Information Processing Systems*, vol. 14, no. 6, pp. 1431-1437, 2018.

[3] H. W. Kim and Y. S. Jeong, "Secure authentication-management human-centric scheme for trusting personal resource information on mobile cloud computing with blockchain," *Human-centric Computing and Information Sciences*, vol. 8, article no. 11, 2018.

[4] C. C. Fung and N. Jerrat, "A neural network based intelligent intruders detection and tracking system using CCTV images," in *Proceedings of TENCON: Intelligent Systems and Technologies for the New Millennium (Cat. No. 00CH37119)*, Kuala Lumpur, Malaysia, 2000, pp. 409-414.

[5] J. Fisher and M. H. Sanchez, "Authentication and verification of digital data utilizing blockchain technology, "U.S. Patent Application No. 15/083,238, 2016.

[6] D. Gafurov, K. Helkala, and T. Sondrol, "Biometric gait authentication using accelerometer sensor," *Journal of Computers*, vol. 1, no. 7, pp. 51-59, 2006.

[7] M. L. Das, "Two-factor user authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1086-1090, 2009.

[8] J. Zhang, X. Tan, X. Wang, A. Yan, and Z. Qin, "T2FA: transparent two-factor authentication," *IEEE Access*, vol. 6, pp. 32677-32686, 2018.

[9] M. K. Khan and K. Alghathbar, "Cryptanalysis and security improvements of 'two-factor user authentication in wireless sensor networks'," *Sensors*, vol. 10, no. 3, pp. 2450-2459, 2010.

[10] F. Aloul, S. Zahidi, and W. El-Hajj, "Two factor authentication using mobile phones," in *Proceedings of 2009 IEEE/ACS International Conference on Computer Systems and Applications*, Rabat, Morocco, 2009, pp. 641-644.

[11] N. Dadashi, A. W. Stedmon, and T. P. Pridmore, "Semi-automated CCTV surveillance: the effects of system confidence, system accuracy and task complexity on operator vigilance, reliance and workload," *Applied Ergonomics*, vol. 44, no. 5, pp. 730-738, 2013.

[12] "The Verification Code scam in Telecommunication Fraud," 2016 [Online]. Available: http://tv.cctv.com/2016/04/26/VIDE6W0VUOnzLvWsYeLrMDYC160426.shtml.

[13] S. Kumari and M. K. Khan, "More secure smart card-based remote user password authentication scheme with user anonymity," *Security and Communication Networks*, vol. 7, no. 11, pp. 2039-2053, 2014.

[14] Z. Siddiqui, A. H. Abdullah, M. K. Khan, and A. S. Alghamdi, "Smart environment as a service: three factor cloud based user authentication for telecare medical information system," *Journal of Medical Systems*, vol. 38, article no. 9997, 2014.

[15] A. Siddiqui, "Authentication vs Authorization," 2018 [Online]. Available: https://medium.com/datadriven investor/authentication-vs-authorization-716fea914d55.

[16] P. Wang, W. H. Lin, K. M. Chao, and C. C. Lo, "A face-recognition approach using deep reinforcement learning approach for user authentication," in *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)*, Shanghai, China, 2017, pp. 183-188.

[17] A. Braeken, "PUF based authentication protocol for IoT," *Symmetry*, vol. 10, article no. 352, 2018.

[18] P. K. Sharma and J. H. Park, "Blockchain based hybrid network architecture for the smart city," *Future Generation Computer Systems*, vol. 86, pp. 650-655, 2018.

[19] J. R. Agustina and G. G. Clavell, "The impact of CCTV on fundamental rights and crime prevention strategies: the case of the Catalan Control Commission of Video surveillance Devices," *Computer Law & Security Review*, vol. 27, no. 2, pp. 168-174, 2011.

[20] X. Wang, H. Xue, X. Liu, and Q. Pei, "A privacy-preserving edge computation-based face verification system for user authentication," *IEEE Access*, vol. 7, pp. 14186-14197, 2019.

[21] W. H. Lin, P. Wang, and C. F. Tsai, "Face recognition using support vector model classifier for user authentication," *Electronic Commerce Research and Applications*, vol. 18, pp. 71-82, 2016.

[22] P. Samangouei, V. M. Patel, and R. Chellappa, "Facial attributes for active authentication on mobile devices," *Image and Vision Computing*, vol. 58, pp. 181-192, 2017.

[23] M. Sajjad, S. Khan, T. Hussain, K. Muhammad, A. K. Sangaiah, A. Castiglione, C. Esposito, and S. W. Baik, "CNN-based anti-spoofing two-tier multi-factor authentication system," *Pattern Recognition Letters*, 2018. https://doi.org/10.1016/j.patrec.2018.02.015.

[24] C. E. Lee, L. Zheng, Y. Zhang, V. L. Thing, and Y. Y. Chu, "Towards building a remote anti-spoofing face authentication system," in *Proceedings of TENCON 2018: 2018 IEEE Region 10 Conference*, Jeju, Korea, 2018, pp. 0321-0326.

[25] M. Azimpourkivi, U. Topkara, and B. Carbunar, "Camera based two factor authentication through mobile and wearable devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, article no. 35, 2017.

[26] T. Zhu, Z. Qu, H. Xu, J. Zhang, Z. Shao, Y. Chen, S. Prabhakar, and J. Yang, "RiskCog: unobtrusive real-time user authentication on mobile devices in the wild," *IEEE Transactions on Mobile Computing*, 2019. http://doi.org/10.1109/TMC.2019.2892440.

[27] R. Zhang, Y. Xiao, S. Sun, and H. Ma, "Efficient multi-factor authenticated key exchange scheme for mobile communications," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 625-634, 2017.

[28] A. Ometov, V. Petrov, S. Bezzateev, S. Andreev, Y. Koucheryavy, and M. Gerla, "Challenges of multi-factor authentication for securing advanced IoT applications," *IEEE Network*, vol. 33, no. 2, pp. 82-88, 2019.

Byoung-WooK Kwon https://orcid.org/0000-0002-7730-5713

He received B.S. in Department of Computer Science and Engineering, Dongseo University in 2017. Since March 2017, he is with the Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech) as Master Course.

Pradip Kumar Sharma https://orcid.org/0000-0001-6620-9083

He is a research scholar at the Seoul National University of Science and Technology. He works in the Ubiquitous Computing & Security Research Group. Prior to beginning the PhD program, he worked as a software engineer at MAQ Software, India. He worked on a variety of projects, proficient in building large-scale complex data warehouses, OLAP models and reporting solutions that meet business objectives and align IT with business. He received his Master's degree in Computer Science from the Thapar University, in 2014, India. His current research interests are focused on the areas of ubiquitous computing and security, cloud computing, SDN, SNS, and IoT. He is also reviewer top cited journals such as *IEEE Com. Mag.*, *IEEE Net. Mag.*, *IEEE SJ*, *IEEE TII*, *IEEE IoT*, *IEEE TNSM*, *FGCS*, and *IEEE CE Mag.*

James J. (Jong Hyuk) Park https://orcid.org/0000-0003-1831-0309

He received Ph.D. degrees in Graduate School of Information Security from Korea University, Korea and Graduate School of Human Sciences from Waseda University, Japan. From December, 2002 to July, 2007, Dr. Park had been a research scientist of R&D Institute, Hanwha S&C Co. Ltd., Korea. From September, 2007 to August, 2009, He had been a professor at the Department of Computer Science and Engineering, Kyungnam University, Korea. He is now a professor at the Department of Computer Science and Engineering and Department of Interdisciplinary Bio IT Materials, Seoul National University of Science and Technology (SeoulTech), Korea. Dr. Park has published about 200 research papers in international journals and conferences. He has been serving as chair, program committee, or organizing committee chair for many international conferences and workshops. He is a steering chair of international conferences—MUE, FutureTech, CSA, CUTE, UCAWSN, World IT Congress-Jeju. He is editor-in-chief of *Human-centric Computing and Information Sciences* (HCIS) by Springer, The *Journal of Information Processing Systems* (JIPS) by KIPS, and *Journal of Convergence* (JoC) by KIPS CSWRG. He is Associate Editor/Editor of 14 international journals including JoS, JNCA, SCN, CJ, and so on. In addition, he has been serving as a Guest Editor for international journals by some publishers: Springer, Elsevier, John Wiley, Oxford University Press, Emerald, Inderscience, MDPI. He got the best paper awards from ISA-08 and ITCS-11 conferences and the outstanding leadership awards from IEEE HPCC-09, ICA3PP-10, IEE ISPA-11, PDCAT-11, IEEE AINA-15. Furthermore, he got the outstanding research awards from the SeoulTech, 2014. His research interests include IoT, human-centric ubiquitous computing, information security, digital forensics, vehicular cloud computing, multimedia computing, etc. He is a member of the IEEE, IEEE Computer Society, KIPS, and KMMS.