RFaceID: Towards RFID-based Facial Recognition

CHENGWEN LUO, Shenzhen University, China
ZHONGRU YANG, Shenzhen University, China
XINGYU FENG, Shenzhen University, China
JIN ZHANG, Shenzhen University, China
HONG JIA, University of New South Wales, Australia
JIANQIANG LI*, Shenzhen University, China
JIAWEI WU, Shenzhen University, China
WEN HU, University of New South Wales, Australia

Face recognition (FR) has been widely used in many areas nowadays. However, the existing mainstream vision-based facial recognition has limitations such as vulnerability to spoofing attacks, sensitivity to lighting conditions, and high risk of privacy leakage, etc. To address these problems, in this paper we take a sparkly different approach and propose *RFaceID*, a novel RFID-based face recognition system. *RFaceID* only needs the users to shake their faces in front of the RFID tag matrix for a few seconds to get their faces recognized. Through theoretical analysis and experiment validations, the feasibility of the RFID-based face recognition is studied. Multiple data processing and data augmentation techniques are proposed to minimize the negative impact of environmental noises and user dynamics. A deep neural network (DNN) model is designed to characterize both the spatial and temporal feature of face shaking events. We implement the system and extensive evaluation results show that *RFaceID* achieves a high face recognition accuracy at 93.1% for 100 users, which shows the potential of *RFaceID* for future facial recognition applications.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

Additional Key Words and Phrases: RFID, face recognition, data augmentation, neural network

ACM Reference Format:
Chengwen Luo, Zhongru Yang, Xingyu Feng, Jin Zhang, Hong Jia, Jianqiang Li, Jiawei Wu, and Wen Hu. 2021. RFaceID: Towards RFID-based Facial Recognition . *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 170 (December 2021), 21 pages. https://doi.org/10.1145/3494985

1 INTRODUCTION

Due to its wide applications in areas such as public security [1], digital payment [23], device unlocking [2], etc., face recognition has become one of the most researched topics in the literature in recent years [19]. In common

*Corresponding author.

Authors' addresses: Chengwen Luo, chengwen@szu.edu.cn, Shenzhen University, Shenzhen, China; Zhongru Yang, Shenzhen University, Shenzhen, China, 1910273031@email.szu.edu.cn; Xingyu Feng, Shenzhen University, Shenzhen, China, 2172272057@email.szu.edu.cn; Jin Zhang, Shenzhen University, Shenzhen, China, jin.zhang@szu.edu.cn; Hong Jia, University of New South Wales, Sydney, Australia, h.jia@unsw.edu.au; Jianqiang Li, Shenzhen University, Shenzhen, China, lijq@szu.edu.cn; Jiawei Wu, Shenzhen University, Shenzhen, China, gary036@163.com; Wen Hu, University of New South Wales, Sydney, Australia, wen.hu@unsw.edu.au.

2474-9567/2021/12-ART170 $15.00
https://doi.org/10.1145/3494985

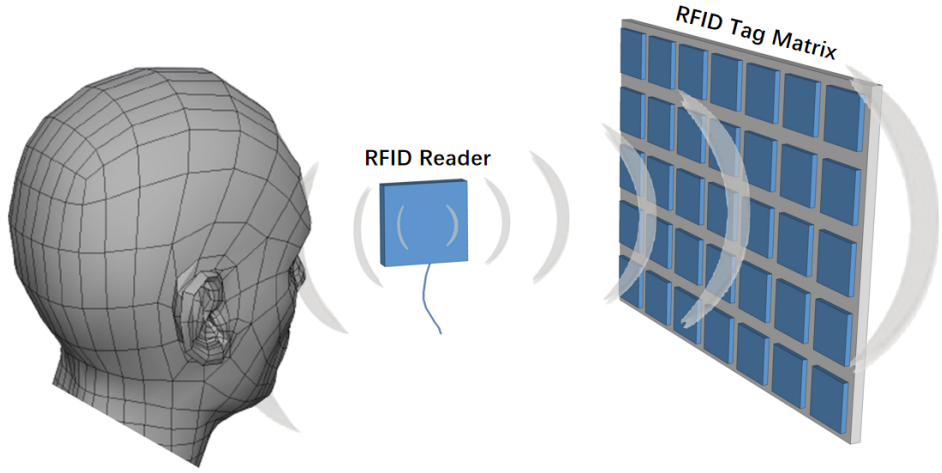


Fig. 1. Rationale of RFID-based face recognition: unique 3D geometry of human face leads to different multi-path reflections.

face recognition scenarios, the acquisition of images and videos requires visual input using cameras. However, the use of cameras for face recognition suffers from several limitations. For example, except for security monitoring demands, common cameras (e.g., on computers and smart devices) are not equipped with infrared lights for the night vision, resulting in lower recognition accuracy in poor lighting conditions. Moreover, the risk of privacy leakage using cameras also hinders the usage of vision-based face recognition systems[17]. Recently, researchers have also discovered that such vision-based face recognition systems make it easy for attackers to obtain the users' face information and extract relevant information for spoofing attacks (including 2D, 3D) [3].

Motivated by the above limitations and the recent developments in wireless sensing technologies [16], we take a starkly different perspective and use RFID signals for face recognition in this paper. The use of RFID signals to realize the perception of the environment and the human body has become a vital application in the field of Internet of Things (IoT) due to the unobtrusive nature of the wireless sensing applications. For example, in [10, 22], human body positioning and trajectory tracking are perceived through RFID signals. [6, 7, 15] realize the activity recognition of single or even multiple people through RFID signals. All these work has shown the feasibility of RFID-based fine-grained human sensing and the potential of using RFID signals for human face recognition as well. The idea of RFID-based face recognition is illustrated in Fig. 1. A tag matrix composed of RFID tags is placed in front of the human face, meanwhile the time series of received signal strength (RSS) and phase are collected. In theory, the 3D geometry and inner biomaterial feature of different faces are different which leads to different multi-path reflections produced by the face, and such differences in RFID signals might provide good opportunities for efficient RFID-based face recognition.

Inspired by the above idea, in this paper we further promote the idea of wireless sensing and propose a novel wireless sensing system to realize the RFID-based face recognition. Such system has several advantages. First, *privacy-preserving*. Since no image or video is captured for face recognition, such system significantly reduces the risk of privacy leakage. Second, *robust to different lighting conditions.* The wireless signal is not affected by the change of lighting conditions and the system should work in complete darkness. Third, *more reliable against spoofing attacks.* Recent study has revealed that radio frequency (RF) signals are sensitive to the material they reflect during propagation, this makes the wireless-based face recognition more robust to spoofing attacks even attackers are able to manipulate 3D-printed masks [21]. As such, the RFID-based face recognition technology

has the potential to become one important complement to the current popular vision-based face recognition technologies.

However, it is still not trivial to achieve robust and accurate face recognition using RFID signals and several challenges need to be tackled: (1) *Environmental noise.* One important factor that affects the robustness of wireless sensing systems is the multi-path effect. The multi-path signals contributed by objects such as the moving people in the background will be captured by the RFID reader as noises and affect the final recognition accuracy. Environmental noises need to be properly addressed to achieve a robust face recognition performance. (2) *Spatial-temporal pattern recognition.* The face recognition process not only involves the spatial pattern recognition (i.e., face 3D geometry), but also involves the temporal patterns (i.e., shaking patterns of face), therefore a neural network structure need to be properly designed to capture both the spatial and temporal features in the face recognition process. (3) *Small-scale training data set.* Furthermore, in practice it is usually hard to obtain sufficient training data for each user. As a result, with the lack of sufficient training data in most practical scenarios, the over-fitting problem must be tackled under the small-scale training data set. (4) *Distance independence.* Finally, due to the physical characteristics of the wireless signals, the distance between the face and the tag matrix has direct impact on the phase and RSS patterns captured by the RFID reader, hence affecting the final recognition performance. Achieving distance independence is non-trivial due to the unknown geometry of arbitrary face, and the system needs to be specially designed to alleviate the effect of changing distances.

To address the above challenges, in this paper we propose *RFaceID*, a novel RFID-based face recognition system. *RFaceID* incorporates various signal processing techniques and artificial intelligence techniques to achieve robust and accurate face recognition using RFID wireless signals. A novel Deep Neural Network (DNN) model composed of Convolutional Neural Network (CNN) [14] and the bi-directional Long Short-Term Memory (Bi-LSTM) neural network [18] is proposed to capture both the spatial and temporal characteristics of human face. To alleviate the environmental noises and user dynamics, conquer the small-scale training data problem, and to improve the overall robustness, various data augmentation techniques are proposed. Finally, extensive evaluations with 100 subjects are conducted which show that *RFaceID* achieves a high classification accuracy at 93.1%. Security analysis and user studies also show the usability of the system in practice.

The contributions of *RFaceID* are summarized as follows:

- We propose *RFaceID*, a novel face recognition system based on Commercial-Off-The-Shelf (COST) RFID, which is device-free, non-obtrusive, not sensitive to lighting conditions and minimizes the risk of privacy leakage.
- We design an effective neural network model incorporating the spatial pattern recognition capability of CNN, and temporal pattern recognition capability of LSTM to effectively recognize the unique features of user face during the dynamic recognition process.
- We propose a set of novel data augmentation techniques, which are able to effectively compensate for the negative impact of environmental noises, face shaking speed variations, face shaking direction variations, distance variations, etc., and address the over-fitting problem caused by the small-scale training data.

The rest of the article is organized as follows. Section 2 introduces the relevant technical background and discusses the feasibility of using RFID signals for face recognition. Section 3 provides the detailed design of *RFaceID*. Section 4 evaluates the overall system performance, and Section 5 discusses the limitations and future work of the system. Finally Section 6 concludes our work.

2 TECHNICAL BACKGROUND

2.1 Theoretical Feasibility of RFID-based Face Recognition

In this section, we will discuss in detail the multi-path effect of the face on the RFID signal, so as to theoretically analyze the feasibility of RFID-based face recognition. Common RFID system consists of readers, antennas and

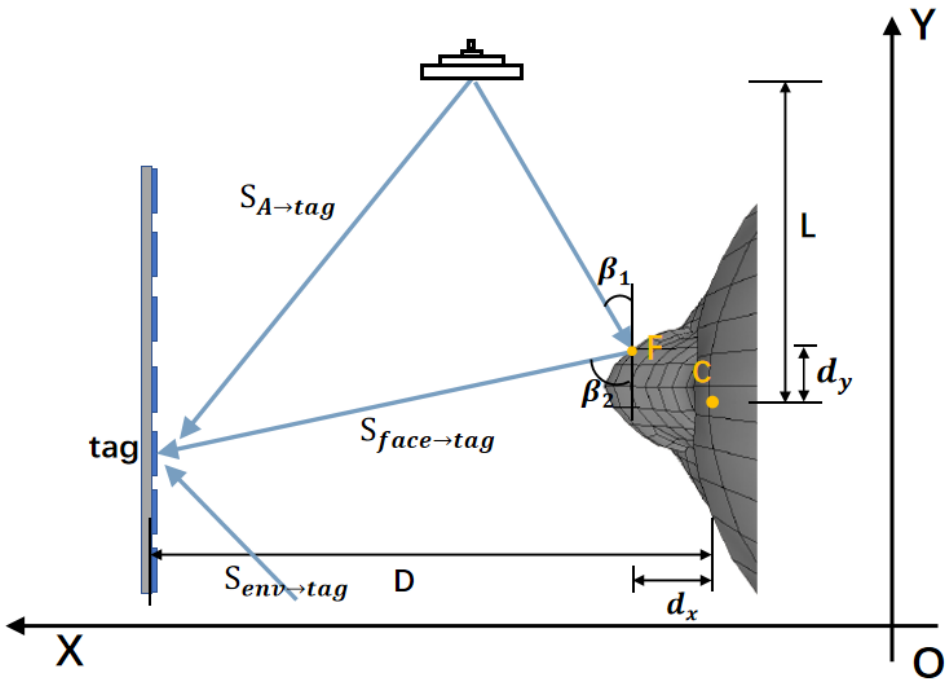


Fig. 2. multi-path effect of a human face in front of the tag matrix

electronic tags. The communication and energy supply between the reader and the electronic tag are completed by electromagnetic backscatter coupling. The passive electronic tag captures the radio frequency signal emitted by the reader for energy, and then modulates its Electronic Product Code (EPC) into the backscatter signal through ON-OFF, which will realize the information exchange with the reader [5]. The reader obtains indicators (such as phase and RSS) from the backscatter signal, and the amplitude of these indicators is affected by the distance between the tag and the antenna and the environmental multi-path signals.

The multi-path effect of human face is illustrated in Fig.2. As shown in Fig.2,the distance between the center point C of the face and the antenna and tag matrix is denoted as L and D, i.e., the vertical distance and horizontal distance respectively. We assume that RFID signal S is reflected by arbitrary point F on the face to a certain tag. The distances between point F and the antenna and tag matrix are denoted as $L - d_y$ and $D - d_x$ respectively. Here d_y and d_x represent the unique 3D geometry of the face.

The relation between the RFID signal S and the phase and RSS can be denoted as:

$$S \approx \alpha \cdot e^{j \cdot \theta} \tag{1}$$

where θ is the phase of RFID signal, and the amplitude α can be directlty coverted to the RSS value. The attenuation of signal propagation satisfies the following equation:

$$S' = h \cdot S\,, \tag{2}$$

where S is the source signal, S' is the attenuated signal, and h represents the attenuation coefficient in the signal propagation process, which relates to the propagation distance of the signal, the angle of reflection, and the material of the reflector. The relationship between the attenuation coefficient h and the propagation distance d

can be formulated as:

$$h = \frac{1}{d^2} \cdot e^{j \cdot \theta} , \tag{3}$$

and θ can be formulated as:

$$\theta = (2\pi \cdot \frac{d}{\lambda}) \ mod \ 2\pi , \tag{4}$$

where λ is the wavelength of the radio frequency signal and is a constant.

As shown Fig.2, the RF signal captured by the tag comes from 3 parts, directly from the RF signal $S_{A \to tag}$ emitted by the antenna, the signal $S_{face \to tag}$ reflected by the face, and the reflected signal $S_{env \to tag}$ from other surrounding objects. The total signal (marked as S_{tag}) obtained by the tag therefore can be expressed as:

$$S_{tag} = S_{A \to tag} + S_{face \to tag} + S_{env \to tag} , \tag{5}$$

Since the distance between the tag and the antenna is fixed, $S_{A \to tag}$ can be regarded as a constant. Although the surrounding environment is changing (such as the movement of objects and people), the human face is much closer to the label matrix compared to the surrounding environment, and the reflection signal $S_{face \to tag}$ of the human face is much larger than the reflected signal $S_{env \to tag}$ of the environment, so we only consider the multi-path influence of the face on the label, and treat $S_{env \to tag}$ as a constant.

The multi-path signal of the face has gone through three stages: the source signal propagates from the antenna to the face, then the signal is reflected and refracted on the face surface, and finally the signal propagates from the reflection of the face to the tag, $S_{face \to tag}$ then can be calculated as:

$$S_{face \to tag} = S \cdot h_{A \to face} \cdot h_{face} \cdot h_{face \to tag} . \tag{6}$$

The $h_{A \to face}$ represents the signal attenuation coefficient when the signal propagates from the antenna to the face, which is related to the distance $d_{A \to face}$. Figure 2 shows that $d_{A \to face}$ can be calculated as:

$$d_{A \to face} = \frac{L - d_y}{cos\beta_1} , \tag{7}$$

According to Eq.(3), Eq.(4) and Eq.(7), it can be expressed as:

$$h_{A \to face} = \frac{{cos\beta_1}^2}{(L - d_y)^2} \cdot e^{j \cdot \theta_{A \to face}} , \tag{8}$$

where $\theta_{A \to face}$ can be formulated as:

$$\theta_{A \to face} = (2\pi \cdot \frac{L - d_y}{\lambda \cdot cos\beta_1}) \ mod \ 2\pi , \tag{9}$$

The $h_{face \to tag}$ represents the signal attenuation coefficient when the signal propagates from the face to the tag, which is related to the distance $d_{face \to tag}$ between the two. Figure 2 shows that $d_{face \to tag}$ can be calculated as:

$$d_{face \to tag} = \frac{D - d_x}{sin\beta_2} , \tag{10}$$

According to Eq.(3), Eq.(4) and Eq.(10),the $h_{face \to tag}$ can be expressed as:

$$h_{face \to tag} = \frac{{sin\beta_2}^2}{(D - d_x)^2} \cdot e^{j \cdot \theta_{face \to tag}} , \tag{11}$$

where $\theta_{face \to tag}$ can be formulated as:

$$\theta_{face \to tag} = (2\pi \cdot \frac{D - d_x}{\lambda sin\beta_2}) \ mod \ 2\pi , \tag{12}$$

The h_{face} represents the signal attenuation coefficient when the signal is reflected and refracted on the face. The signal will not only reflect on the face, but also some of the signal will be refracted on the face. The structure of the face can be seen as a multi-layer mixed material. When the signal is refracted into the face, almost all the signal will be captured by the human body, which forms the loss of signal energy and phase change [21]. Thus h_{face} can be represented as:

$$h_{face} = \sqrt{R_{per}} \cdot e^{j \cdot \theta_{per}} , \tag{13}$$

where $\sqrt{R_{per}}$, θ_{per} are related to the geometry and material of the face according to Snell's Law[20]. Finally, the signal S_{last} read by the reader can be expressed as:

$$S_{last} = S_{tag} \cdot h_{tag \to A} , \tag{14}$$

Since the position between the antenna and the tag matrix is fixed, $h_{tag \to A}$ is constant.

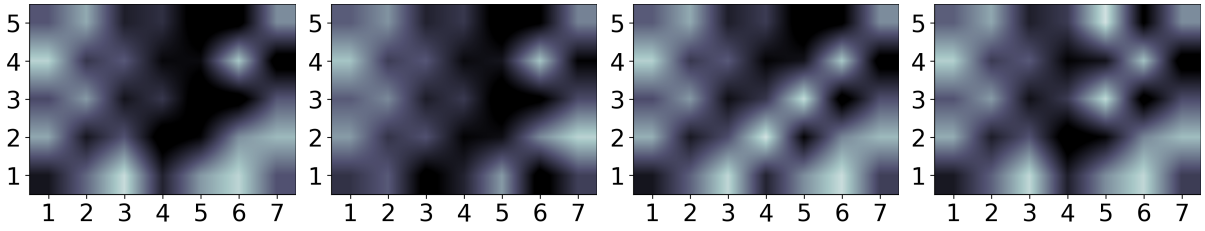
According to Eq.(1,5,6,8,11,13,14), the signal amplitude α related to the RSS which reader receives can be derived as:

$$\alpha = \frac{cos\beta_1{}^2 \cdot sin\beta_2{}^2}{(L - d_y)^2 (D - d_x)^2} \cdot \sqrt{R_{per}} \cdot 2\alpha_{tag \to A} \cdot \alpha_{env} , \tag{15}$$

where $\alpha_{tag \to A}$, α_{env} are constants. L and D are related to the position of the face. And $cos\beta_1$, $sin\beta_2$, d_x, d_y and $\sqrt{R_{per}}$ are related to the 3D geometry and internal materials of the human face. According to Eq.(1,5,6,9,12,13,14), the phase θ_{last} which reader receives can be derived as:

$$\theta_{last} = [2\pi \cdot (\frac{L - d_y}{\lambda \cdot cos\beta_1} + \frac{D - d_x}{\lambda \cdot sin\beta_2} + \theta_{per} + \theta_{env} + 2\theta_{tag \to A})] \ mod \ 2\pi , \tag{16}$$

where θ_{env}, $\theta_{tag \to A}$, λ are constants. L and D are related to the position of the face. And θ_{per}, $cos\beta_1$, $sin\beta_2$, d_x, d_y are related to the 3D geometry and internal materials of the human face. In summary, the phase and RSS collected by the reader under the influence of face multi-path are directly related to the 3D geometry of face, internal materials of the face and the position of the face relative to the antenna and tag matrix. Therefore, theoretically the uniqueness of user faces provides the feasibility for RFID-based face recognition, and this lays the foundation for our system design.

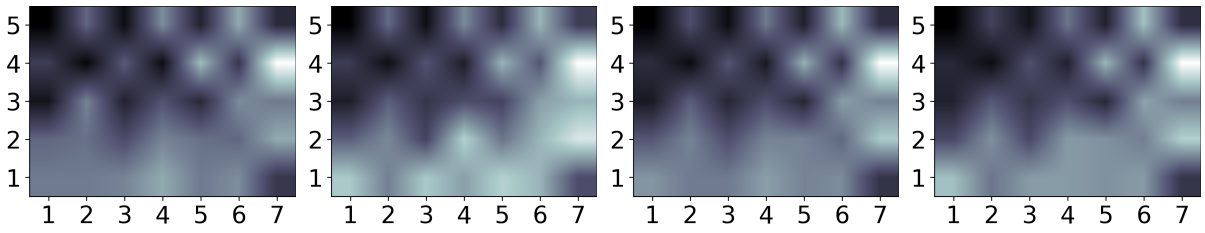


(a) Person A (b) Person B (c) Person C (d) Person C (different angle)

Fig. 3. Differences in phase distributions with different user faces keeping static in front of a 5×7 tag matrix

2.2 Experimental Validation of Feasibility of RFID-based Face Recognition

To experimentally validate the feasibility of RFID-based face recognition, we conduct two different types of preliminary experiments. In the first experiment (Fig. 3 and Fig. 4), three different users are kept static in front of the tag matrix one by one. As shown in Fig. 3a, Fig. 3b, and Fig. 3c, the uniqueness of three different faces results

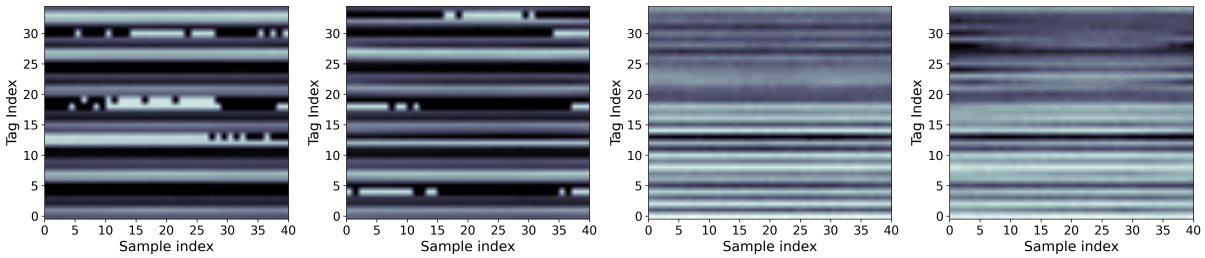


(a) Person A (b) Person B (c) Person C (d) Person C (different angle)

Fig. 4. Differences in RSS distributions with different user faces keeping static in front of a 5×7 tag matrix

in different phase patterns perceived by a 5×7 RFID tag matrix. Similarly, as shown in 4a, Fig. 4b, and Fig. 4c, the RSS values of the tag matrix also present unique patterns for different faces.

However, when the user performs face recognition, the direction of the face cannot always be perpendicular to the tag matrix, there might be angles between the face and the tag matrix. And the difference in the angles between the face and the tag matrix may lead to different phase and RSS patterns percepted by the RFID reader. As an illustration, in Fig.3c and Fig.3d, same user face but different angles also leads to different phase patterns. Similarly, as shown in Fig.4c and Fig.4d, RSS patterns also affected by different angles. As such, in our system design, users are required to shake their faces from left and right during face recognition. This not only reduces the error in the recognition caused by the angle differences between the face and the tag matrix, but also introduce temporal features of the user face which can be utilized to improve the robustness of the system. As shown in Figure 5, the differences in the phase and RSS patterns when users are shaking faces provides sufficient discrimination power to effectively distinguish different faces, making the RFID-based face recognition feasible in practice.



(a) Person A (b) Person B (c) Person A (d) Person B

Fig. 5. Temporal patterns of phase ((a),(b)) and RSS ((c),(d)) when users are shaking their faces in front of the tag matrix

3 RFaceID: DATA PROCESSING AND MODEL CONSTRUCTION

3.1 Trust and Threat Models

We envision the use of *RFaceID* primarily as a complementary approach for conventional facial authentication system or as part of the two-factor authentication systems to authenticate the identity of the user to prevent spoof attacks. *RFaceID* addresses the issue of poor lighting conditions by using radio instead of cameras. In a RFID-based facial authentication system, each user is required to shake their faces in front of a tag matrix for

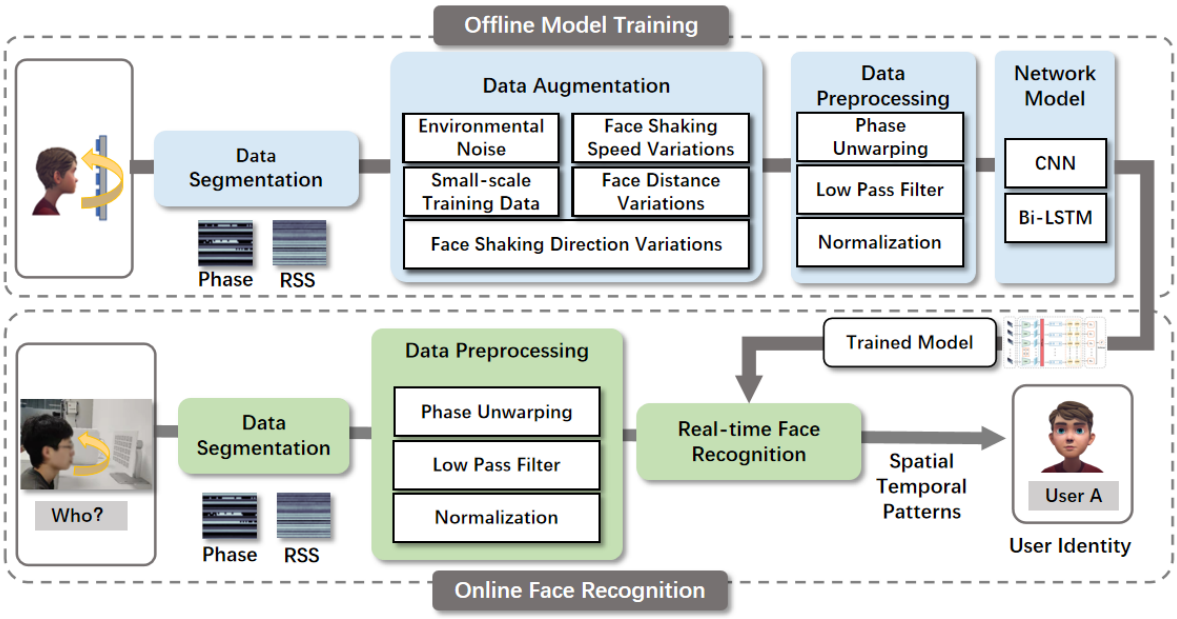


Fig. 6. System overview of *RFaceID*

data collection. The RFID reader first collects face data and transmits them to the sever. The server will then perform authentication to verify the user's identity by using the radio data. In this paper, we assume the RF signal collected by RFID readers devices are trustworthy. Also, our system trusts the communication channel between the RFID readers and the authentication server. How to secure the RFID and server hardware is out of the scope of this paper. Furthermore, we assume that Denial of Service (DoS) or radio jamming attacks is out of consideration in this paper since such attacks are easy to be detected (e.g., by monitoring idle radio signal) and mitigated.

However, the aforementioned authentication system is vulnerable to user spoofing attacks. For instance, an adversary may pretend to be another person to gain access (e.g., to the building). Therefore, the adversary model considered in this paper focuses on impersonation attacks. We assume the presence of two types of impersonation attacks: a passive adversary and an active adversary. The passive adversary tries to spoof the authentication system by using her own face appearance and motions. The active spoofing attacker knows the authentication scheme and will try her best to imitate the appearance and motion of the genuine user to spoof the authentication system.

3.2 System Overview

The system overview of *RFaceID* is shown in Fig. 6. As shown in the figure, the *RFaceID* consists of two different stages: the *offline training* stage and the *online face recognition* stage. In the offline model training stage, the face shaking data of all users is collected. A data segmentation module is firstly used to capture the start and end event of the face shaking activity. Each data segment $\boldsymbol{S} = \{\boldsymbol{\theta}_N, \boldsymbol{RSS}_N\}$, both the phase series $\boldsymbol{\theta}_N = \{\boldsymbol{\theta}_i, i = 1, 2, ..., N\}$ and RSS series $\boldsymbol{RSS}_N = \{\boldsymbol{RSS}_i, i = 1, 2, ..., N\}$ are captured. Here N is the total number of tags in the tag matrix. And each tag i captures a phase series $\boldsymbol{\theta}_i = \{\theta_i^j, j = 1, 2, ..., T\}$ and a RSS series $\boldsymbol{RSS}_i = \{RSS_i^j, j = 1, 2, ..., T\}$ of T samples over time. Each segment of T samples, i.e., $\boldsymbol{\theta}_N$ and $\boldsymbol{RSS}_N$, containing one complete face cycle is captured as one training sample. Since the face recognition accuracy can be affected by multiple factors such as

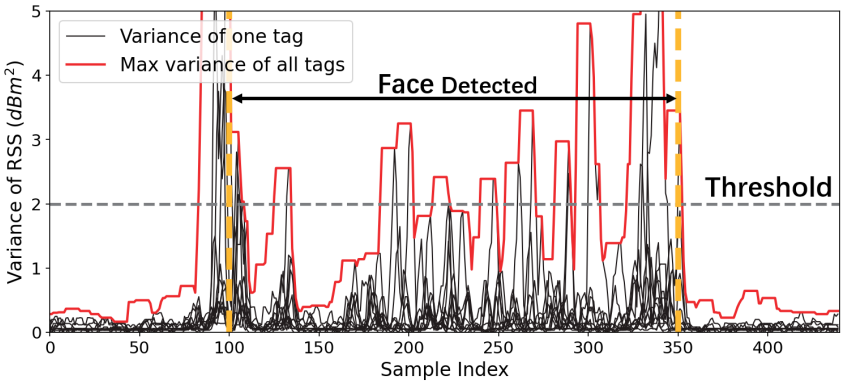


Fig. 7. Detecting face recognition events for data segmentation

over-fitting due to small-scale training data set, environmental noises, distances, etc., in *RFaceID* we introduce a data augmentation module to effectively compensate for the insufficiency of high quality training data and improve the overall robustness of the system. A data preprocessing module is then applied to remove noises in the data and finally a neural network combing the spatial feature extraction capability of CNN and temporal feature extraction capability of Bi-LSTM is designed to train the model for face recognition. In the online face recognition state, the data segmentation and preprocessing modules are applied to the real-time data and the neural network model is used to obtain the identify prediction for each user face.

3.3 Data Segmentation

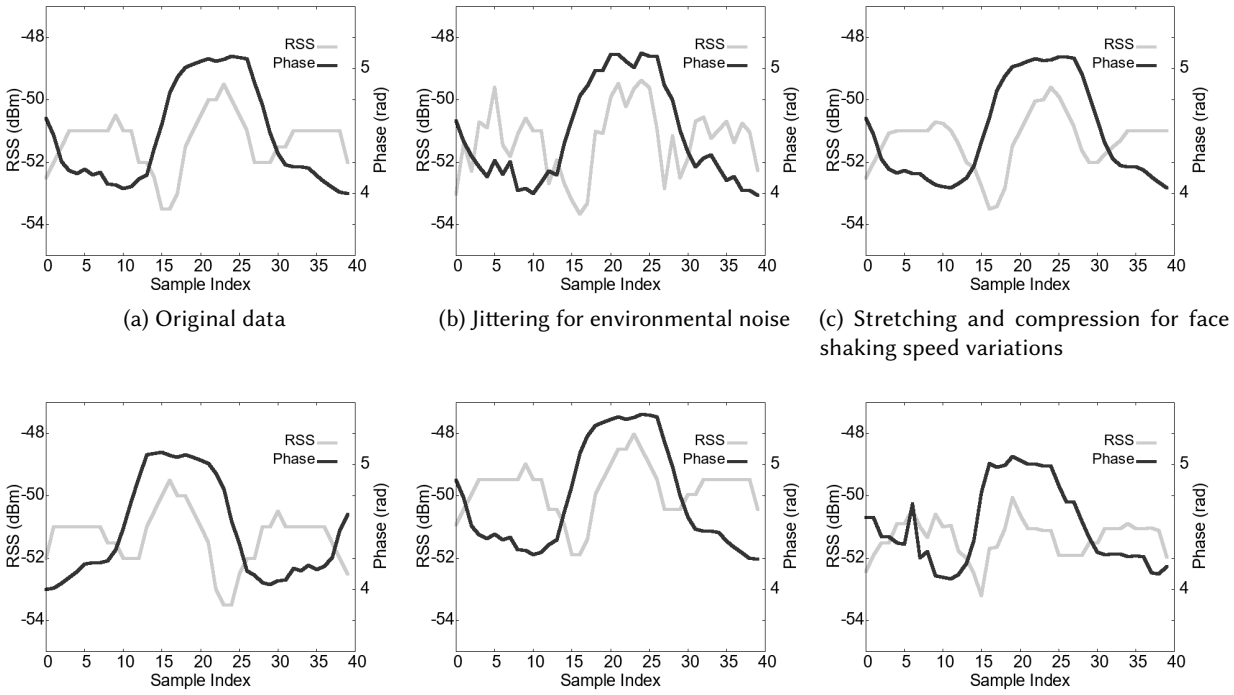
In both the offline training data collection and online face recognition stage, we need to mark the start and end of the continuous phase and RSS sequences to form a segment of face shaking event and get both $\boldsymbol{\theta}_N$ and $\boldsymbol{RSS}_N$. Based on the observations that when a user's face is close to the label matrix for face recognition, the RF-signal collected by the reader will change drastically due to the introduction of multi-path components reflected by the face, in *RFaceID* we use the fluctuation of the RSS as an indicator for data segmentation.

As shown in Fig.7, when a user starts to perform face recognition at time t_1 (sample index 100 in the figure), the RSS values will start to fluctuate correspondingly, which shows that the RSS variance can be used as an efficient indicator for data segmentation. As such, we use a fixed-size sliding variance window to detect the segmentation points. We use a RSS variance threshold th_{RSS} to detect the start and end of the face shaking event. Since different tags in the tag matrix have different RSS variances in the same sliding window, we take the maximum difference of all tags as the current window variance.

3.4 Data Augmentation

In *RFaceID*, the face recognition performance can be affected by multiple factors such as the environmental and user dynamics. The environmental dynamics includes the noises introduced by the changing environment, and the user dynamics includes the variations in face shaking speed, face shaking directions, distances, etc. Besides, the model training also suffers from the small-scale data set and over-fitting problem. In *RFaceID*, we propose several data augmentation techniques to address the above challenges.

3.4.1 Environmental noises. Affected by the environmental noises, phase and RSS readings of RFID tags are usually unstable. Especially the phase is more sensitive to environmental noises. Fortunately, since the face is much closer to tag matrix comparing with other moving objects in the background in our experiment setting illustrated in Fig. 1, the face contributes the dominate part for phase and RSS changes captured by the reader.



(a) Original data (b) Jittering for environmental noise (c) Stretching and compression for face shaking speed variations

(d) Flipping for face shaking direction variations (e) Scaling for distance variations (f) WA-DTW to generate synthesized data

Fig. 8. Illustration of data augmentation methods on phase and RSS

To compensate for the random disturbances introduced by the background objects, we randomly add Gaussian noises with a mean value of zero and a variance of σ to the augmented data set. Since phase and RSS are different in sensitivity to the environmental noises, we set different variances for phase and RSS values. As shown in Fig.8b, jittering is introduced to the phase and RSS data set with variance σ_{phase}=0.05 for phase and σ_{RSS}=0.5 for RSS. The introduced noises models the effect of the random environmental dynamics and compensates for the accuracy loss due to environmental noises.

3.4.2 Face shaking speed variations. When users are performing face recognition, the shaking speed of their faces can be different even for the same user, and the face shaking speed variations will inevitably lead to a degradation in the final recognition accuracy. To address this problem, in *RFaceID* we introduce randomly stretched and compressed phase and RSS series in the augmented data set, which models the effect of different face shaking speed. In the stretching process, due to the continuity of the phase and RSS series, we use interpolation to resample the original time series data to lengthen the original time series data. In the compression process, we perform downsampling operations and generate shorter time series, which corresponds to faster face shaking speeds. In *RFaceID*, each data segment $\boldsymbol{S} = \{\boldsymbol{\theta}_N, \boldsymbol{RSS}_N\}$ is stretched or compressed by a random stretching/compressing factor β, where $\beta \in (-30\%, +30\%)$. Fig.8c shows an example with β=20%. Through this way, the speed variations of user face shaking activities are modeled in the augmented training data.

3.4.3 Face shaking direction variations. When the same user performs face recognition, the direction of face shaking (e.g., from left to right, and from right to left) can also be different. This will result in different phase and RSS patterns and affect the final recognition performance. Fortunately, the reversed face shaking direction results in the reserved pattern of phase and RSS. As a result, the flipping of the entire phase and RSS time series data can effectively compensate for the direction variations. In *RFaceID*, we randomly take 20% in the training set for flipping. Fig.8d shows the effect of data flipping in the augmented data set.

3.4.4 Face distance variations. As revealed by Eq. 15 and Eq. 16, the distance D between the face and the tag matrix has the direct impact on the phase and RSS series captured by the readers. In practice, the distance D might vary depending on the position of the user face during recognition. Based on Eq. 15, the RSS value can be viewed as inverse proportional to D^2, as a result, we augment the data set by adding:

$$\boldsymbol{RSS}'_N = \alpha_{RSS} \cdot \boldsymbol{RSS}_N \tag{17}$$

to the original data set, and here α_{RSS} is the regression factor between D^2 and the RSS value. Similarly, Based on Eq. 16, the phase value can be viewed as proportional to the distance D, and hence we add:

$$\boldsymbol{\theta}'_N = \alpha_\theta \cdot \boldsymbol{\theta}_N \; mod \; 2\pi \tag{18}$$

in the augmented data set, and here α_θ is the regression factor between the phase and the distance D. Fig.8e shows an example of data augmentation for face distance variations.

3.4.5 Small-scale training data. In the training stage, although we can collect large-scale data sets so that the training set can cover different aspects to avoid over-fitting and improve accuracy, collecting large-scale training data is usually cumbersome in practice. As such, in the data augmentation module we finally propose a weighted average dynamic time warping scheme (WA-DTW) to synthesize new data from the original data set. Dynamic Time Warping (DTW) [8] is a technique for signal alignment, and here we use DTW to align multiple phase series and RSS series for N tags.

Suppose we use WA-DTW to synthesize a new phase series from k segments, i.e., $\{\boldsymbol{\theta}_N^1, \boldsymbol{\theta}_N^2, ..., \boldsymbol{\theta}_N^k\}$, random weights $\{\boldsymbol{w_1}, \boldsymbol{w_2}, ..., \boldsymbol{w_k}\}$ are first generated. A the new phase series $\boldsymbol{\theta}''_N$ is generated by solving the following optimization:

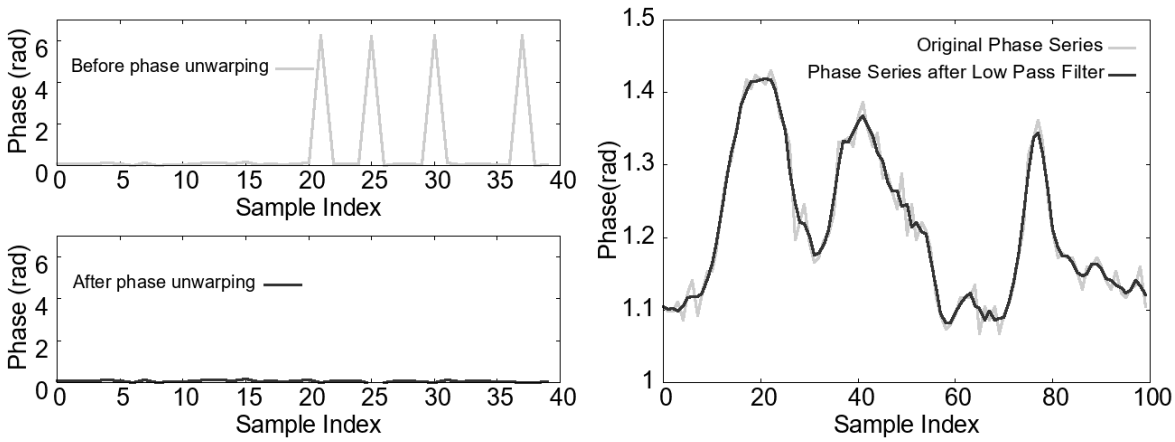
$$argmin_{\boldsymbol{\theta}''_N} \sum_{i=1}^{k} \boldsymbol{w_i} \cdot DTW^2(\boldsymbol{\theta}_N^i, \boldsymbol{\theta}''_N), \tag{19}$$

where $DTW^2(\boldsymbol{\theta}_N^i, \boldsymbol{\theta}''_N)$ finds the DTW distance between the synthesized time series and the original time series, and the weights satisfy $\sum_{i=1}^{k} w_i = 1$. The same process applies to the RSS series to generated weigthed average new synthesized RSS series. As shown in Fig.8f, the new weighted data has a good synthesis of the characteristics of the original data.

3.5 Data Preprocessing

Before passing the augmented data set to modeling training, the data firstly need to be preprocessed. Since the phase value ranges from 0 to 2π, the phase value of some tags varies between 0 and 2π and can fluctuate greatly especially when the phase is close to 0 or 2π, which will inevitably introduce negative impact to the face recognition accuracy. Therefore, we use the Unwarp algorithm [24] to process the original phase and correct the phase value drifting problem. Fig.9a illustrates the effectiveness of the Unwarp algorithm in removing phase noises. Besides, in order to reduce the negative impact of environmental noise on phase and RSS, we use a low pass filter to smooth both the phase and RSS readings. Fig.9b shows the effectiveness of using the low pass filter to remove noises in the data.

In addition, both the phase and RSS values are used for model training, but the orders of magnitude of phase and RSS readings are different. In order to avoid the back-and-forth oscillations and non-convergence during



(a) Phase unwarping (b) Low pass filter

Fig. 9. Data preprocessing

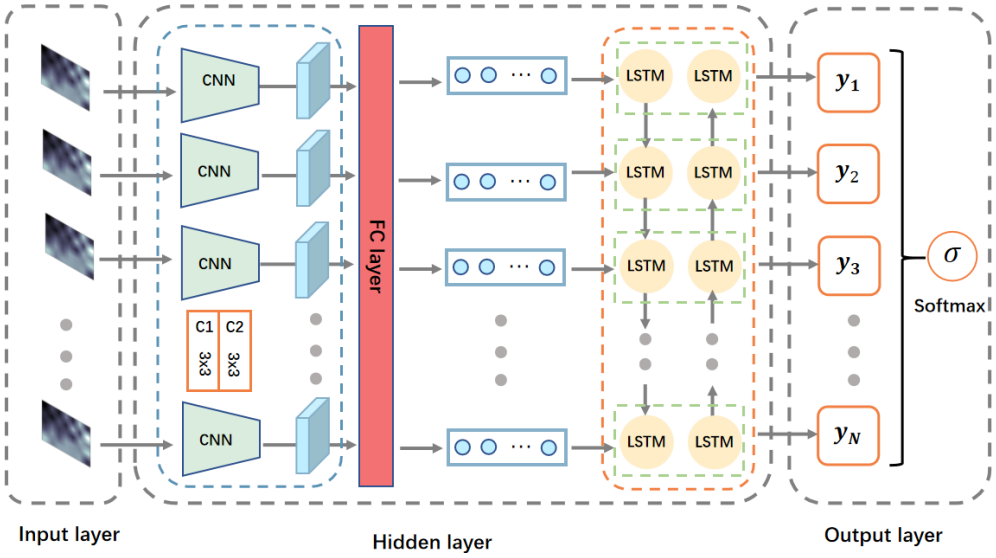


Fig. 10. The neural network architecture used in *RFaceID*

model training, and to improve the convergence speed of the model, feature normalization [12] is also performed in the data preprocessing process.

3.6 Network Model

In *RFaceID*, we use DNN models to capture the the spatial and temporal characteristics of the phase and RSS series. As shown in Figure 10, the neural network in *RFaceID* is composed of three parts: the input layer, hidden layer, and output layer. In the input layer, we use the phase and RSS segments of all tags captured by the RFID reader as input data. Then, a full-connect layer in the hidden layer is used to extract the geometric space and internal material features of the face at different angles at each moment. And the temporal features of face

Input	40 x 70 ⟶ 40 × 5 x 7, 2 channels
Conv	8 hidden, 3 × 3 kernel, preserve shape
ReLU	↓
Conv	16 hidden, 3 × 3 kernel, preserve shape
ReLU	↓
Conv	32 hidden, 3 × 3 kernel, preserve shape
ReLU	↓
Conv	8 hidden, 3 × 3 kernel, preserve shape
ReLU	↓
FC	70 hidden
ReLU	↓
Dropout	p=0.5
BI-LSTM	512 hidden,40 sequence length
Dropout	p=0.5
FC	100 classes
softmax	Negative Log Likelihood

Fig. 11. Architecture of the DNN model used in *RFaceID*

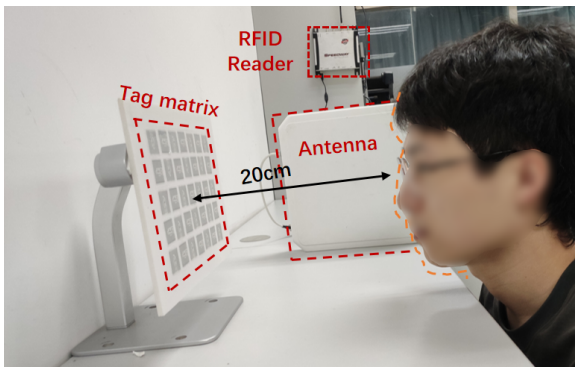


Fig. 12. The experiment setting

shaking is captured through a Bi-LSTM network [18]. Finally, in the output layer, the softmax function is used to multi-classify the features extracted by the LSTM layer to realize face recognition.

Input Layer. The original phase and RSS data will be fed to the input layer as input data after passing through the data preprocessing module. Before the input and output layer, the shape of the data is $T * r * c * 2$. Where T represents the number of samples in a data segment. and $r * c$ represents that the label matrix has r rows and c columns. In our matrix arrangement, $r = 5$ and $c = 7$. And 2 represents the phase and RSS data collected on each tag at the time stamp of each sample.

Hidden Layer. In *RFaceID*, CNN-LSTM neural network is used for multi-classification. The convolution (Conv) layer takes all tag frames as input and outputs continuous time series data, which then forms the input of the Bi-LSTM layer. In this work, CNN extracts the 3D geometry of the face in each tag frame, and then uses a fully connected layer, which uses Rectified Linear Unit (ReLU) as the activation function, to synthesize all the features. Finally, we input the face features in all the tag frames into a Bi-LSTM network to extract the temporal information of the face shaking.

Output Layer. In the output layer, we use a softmax function to regularize the features extracted from the hidden layer and to calculate the probability of each face category y^i when the data is X_j.

$$Pr(y^i \mid X_j) = \frac{e^{X_j^i}}{\sum_{i=1}^{C} e^{X_j^i}} \tag{20}$$

where C represents all categories in the training set.

Our goal is to maximize the estimated probability of all training samples. We use the mean of the cross entropy of all data as the loss function:

$$loss = -\frac{1}{M} \sum_{j=1}^{M} \sum_{i=1}^{N} y^i \cdot Pr(y^i \mid X_j) \tag{21}$$

where M represents the number of all samples in the training set, $y^i \in \{0, 1\}$. And finally we use the Adam algorithm [13] to minimize loss.

4 EVALUATIONS

4.1 Implementation Details and Experimental Settings

In order to verify the effectiveness of the *RFaceID* system in practice, we implemented our system in real environment for evaluation. In this section, we describe the implementation details of the *RFaceID* system.

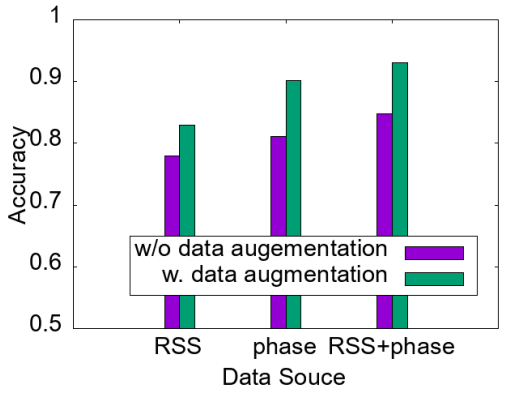


Fig. 13. Impact of data source on recognition accuracy

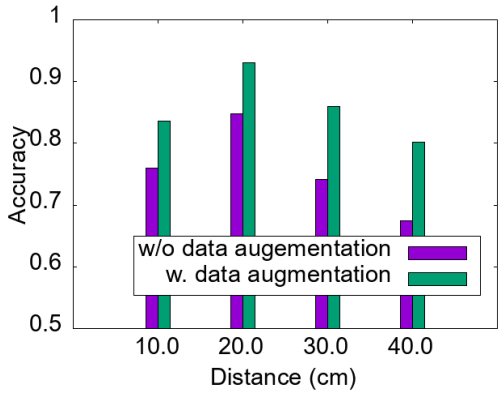


Fig. 14. Impact of distance on recognition accuracy

Hardware: Hardware implementation of *RFaceID* consists of a 5x7 AZ-9629 tag matrix and an Impinj Speedway R420 RFID reader connected to a Laird S9028PCL antenna [11]. The frequency of our fixed antenna is 908.25MHz. This set of equipment can achieve an average sampling frequency of 20Hz for each tag, and the value range of the phase collected by the equipment is $(0, 2\pi)$.

Software: The *RFaceID* consists of two different stages: the offline training stage and the online face recognition stage. In the offline training stage, the neural network model of *RFaceID* is implemented using Tensorflow and runs on server with NVIDIA GeForce RTX 2080 GPU for model training. In the online face recognition stage, we lay out the trained model and complete real-time model inference on a personal computer (PC) with Intel(R) Core(TM) i5-9500 CPU @3.00GHZ and 8 GB RAM. The detail of the DNN model structure is shown in Fig. 11.

Experiment Settings: Fig. 12 shows the deployment of *RFaceID* in the real environment. As shown in the figure, the distance between the center point of the tag matrix and the center point of the face is set to be 20cm, the plane of the tag matrix is perpendicular to the plane of the antenna. In the experiment, there is no strict constraint on the distance, and the distance between the face and the tag matrix can vary depending on the user. To validate the performance of *RFaceID* with sufficient amount of user, we recruit 100 volunteers in total in the experiment. In the evaluation, we use 60% as the training data and the rest as the testing data. To understand the impact of user appearance changes, we ask volunteers to collect data at different times when they wear different clothes, glasses, hats, or face masks. This dataset is used as an additional testing dataset to evaluate the robustness of the system under user appearance changes in Section 4.6. We use classification accuracy [4] as the evaluation metric for *RFaceID*, where:

$$Accuracy = \frac{|TN| + |TP|}{|FN| + |FP| + |TN| + |TP|} \tag{22}$$

and TP, TN, FP, FN represent the true positives, true negatives, false positives, and false negatives, respectively.

4.2 Impact of the Data Source

Phase and RSS data are used in *RFaceID* for face recognition, which have their own different properties. Comparing with phase, RSS has better stability, unlike phase where there is a warp of data (some data will jump between two values of 0 and 2π). However phase also has higher discriminative power and can distinguish subtle differences caused by the differences of human faces. In order to show the classification performance when using different data sources, we use phase, RSS and the combination of two respectively to perform face recognition. As shown in Fig.13, using only RSS or phase, the recognition accuracy of 82.9% and 90.1% are achieved respectively. And the combination of phase and RSS achieves higher classification accuracy at 93.1%. This shows that using the

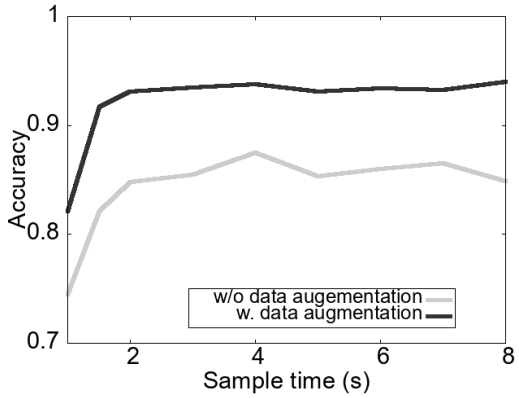


Fig. 15. Impact of sampling time on recognition accuracy

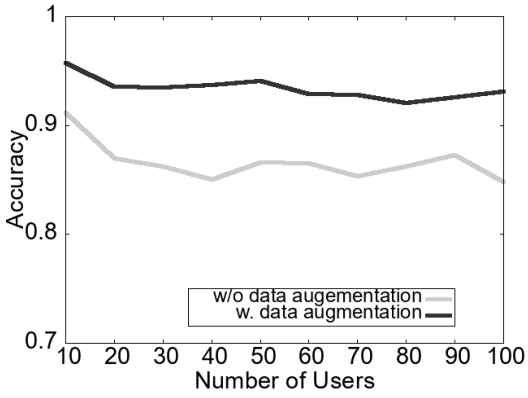


Fig. 16. Impact of user number on recognition accuracy

combination of RSS and phase captures more face-related information, and can achieve better face recognition accuracy.

4.3 Impact of Face-to-Tag Distances

In order to understand the impact of the distance between the face and the tag matrix to the final recognition accuracy, we vary the distances and analyze the corresponding accuracy changes. As shown in Fig.14, we vary the distances from 10 to 40cm, which is sufficient to cover the range of distances when user is performing face recognition in front of the tag matrix. As shown in the figure, the closer the face is to the label matrix, the higher the recognition accuracy except for the distance is at 10cm. As discussed in Section 2 that the greater the distance between the face and the label matrix, the greater the attenuation during signal propagation. This causes the multi-path effect of the human face to have a smaller impact on the tag matrix. When the distance is too large, the multi-path effect of the human face will be masked by the environmental noise. But the face should also not be too close to the tag matrix, or it will result the head to be between the tag matrix and the antenna, resulting a large number of tags cannot be read by the reader and a degradation in the classification accuracy. At 20cm, the system obtains the best classification accuracy at 84.8% without data augmentation and 93.1% with data augmentation. However, even the distance increase from 20cm to 40cm, we still have around 80.2% accuracy.

4.4 Impact of Data Segment Length

In *RFaceID*, although the data segmentation is automatically achieved by the data segmentation module, same data segment length is required for the purpose of model training and real-time classification. As a result, we keep a fix segment length for each data segment in the experiment by discarding the data outside the time window for the overlength segments and interpolating for the shorter segments. To evaluate the impact of the segment length, we vary the segment length while keeping all other settings the same. Fig.15 show that when the segment length is not less than 2s, there is a high and stable classification accuracy at about 93.1%, but when the time is less than 2s, the classification accuracy has a significant drop. This is because volunteers usually complete a round of face shaking in about 2 seconds, and shorter segments are not able to capture the whole cycle of face shaking event. Based on the observation, we adopt 2s as the segment length to balance the recognition speed and the accuracy.

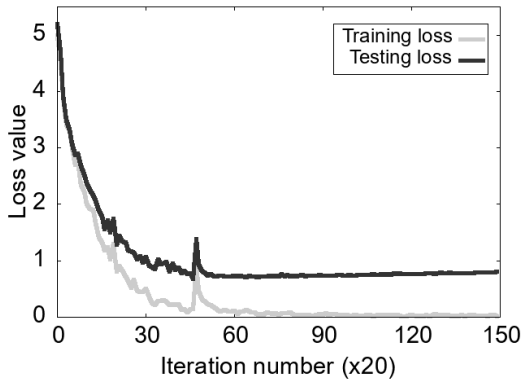


Fig. 17. Learning curve without data augmentation

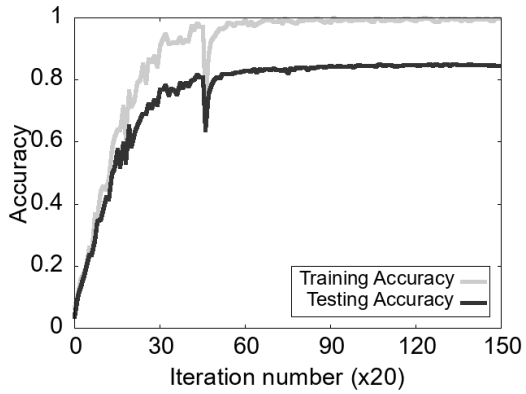


Fig. 18. Accuracy curve without data augmentation

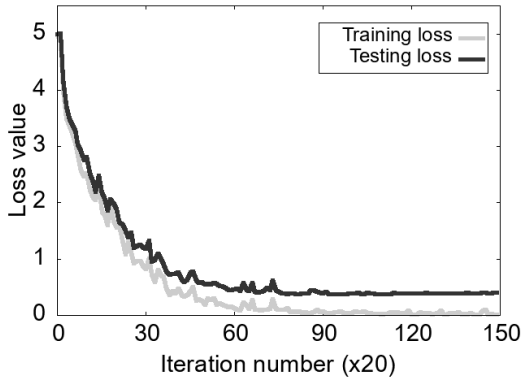


Fig. 19. Learning curve with data augmentation

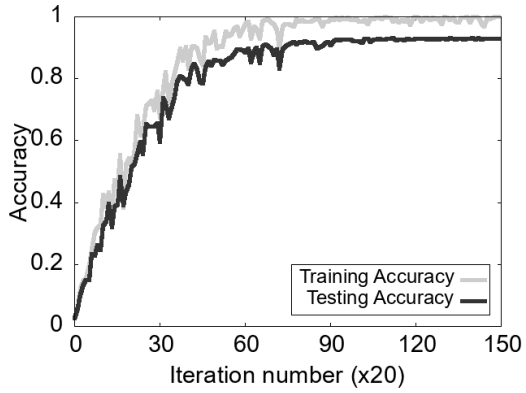


Fig. 20. Accuracy curve with data augmentation

4.5 Impact of User Numbers

In order to further understand the impact of the user number and the stability of *RFaceID*, we conduct experiments to study the recognition accuracy of the system when identifying different numbers of users. We change the number of volunteers from 10 to 100 and analyze the recognition accuracy of the system. As shown in the Fig.16, when the user number is 10 the classification accuracy is 95.7%. Despite the increase in the number of users identified, the *RFaceID* still maintains a high recognition accuracy at 93.1% when the user number reaches 100, which shows that our scheme has a good scalability when increasing the number of users in the system.

4.6 Impact of User Appearance Changes

To evaluate the impact of changes in facial appearance and occlusion on classification accuracy, we conduct experiments to study the effect of wearing different clothes, glasses, hats, and face masks. In this experiment, the users are asked to conduct facial recognition at different times when they change their facial appearances while using the original DNN model for recognition. As shown in the Fig.21, comparing with the original recognition accuracy of 93.1%, wearing different clothes and wearing glasses still achieve similar recognition accuracy at 92.5% and 90.2% respectively. This shows that the usability of *RFaceID* remains high when users change clothes and glasses. However, when the volunteers wear a hat during the test, the average recognition accuracy drops to 82%, And when the volunteers wear a face mask, the accuracy is severely affected and reduces to 59%. This result

Fig. 21. Impact of changing user appearance on the recognition accuracy

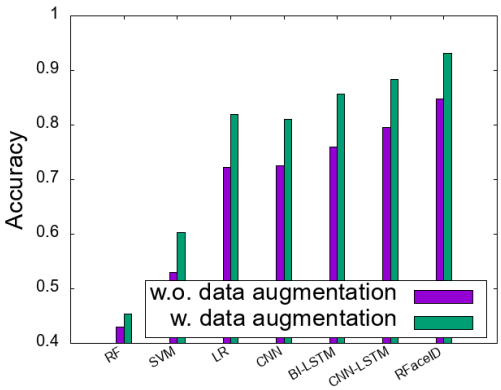


Fig. 22. Impact of different machine learning models

coincides with the fact that *RFaceID* recognizes the 3D geometry of users' faces, and occlusion of the volunteers' face directly affects the recognition accuracy.

4.7 Final Classification Performance

In *RFaceID*, the data augmentation schemes are proposed to alleviate the model over-fitting problem caused by small data sets and other factors such as environmental noises and user dynamics. The learning curve in Fig.17 and accuracy curve in Fig. 18 show that although the highest accuracy of 99.6% can be achieved on the training set, only 84.8% of the recognition accuracy can be achieved on the testing data set in the face recognition stage, which shows that without data augmentation, the model suffers from the over-fitting problem. As a comparison, the learning curve in Fig. 19 and accuracy curve in Fig. 20 show that when we adopt the proposed data augmentation scheme, not only can we achieve 100.0% accuracy on the training set, but also 93.1% accuracy on the testing set. It shows that the data augmentation scheme proposed can effectively alleviate the over-fitting problem caused by the small-scale data set. Fig. 22 shows the comparison among different machine learning models used. As shown in the figure, the CNN+Bi-LSTM model adopted in *RFaceID* achieves the highest accuracy comparing with Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), CNN only, Bi-LSTM only, and CNN+LSTM. As shown in Fig.22, in *RFaceID* all 100 users can be efficiently recognized using the phase and RSS information of RFID, with an average of 93.1% face recognition accuracy.

4.8 Robustness against Attackers

As discussed in Section 3.1 above, we assume the presence of a passive adversary and an active attacker during an authentication session. We evaluate the robustness of the proposed system against the eavesdropper and active attacker by conducting the following two imposter attempt experiments.

- A passive imposter attempt is an attempt when an imposter performs authentication using her face appearance and motion. This attack happens when the attacker appears in front of the *RFaceID* tag matrix and attempt to gain access.
- An active imposter attempt means the imposter mimics the appearance and motion of one specific genuine user with the aim to spoof the authentication system. This attack happens when the attacker mimic the facial appearance and motion of a genuine user front of the *RFaceID* tag matrix and attempt to gain access.

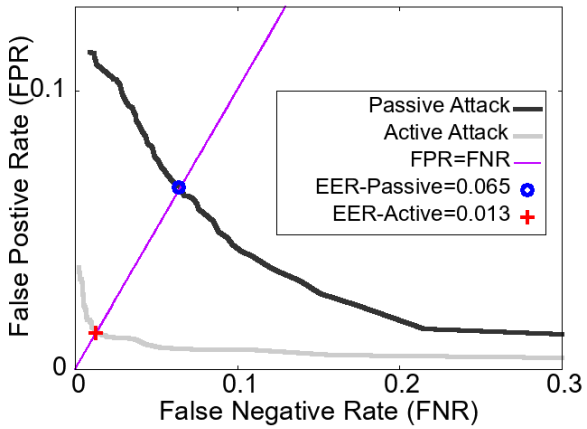


Fig. 23. DET curve of *RFaceID* under passive and active attacks

The first experiment is conducted to evaluate the robustness to a passive imposter. In this experiment, we use the raw radio signal from each subject of all 100 subjects from the dataset as passive imposter attempts and test if the imposter can be authenticated as other users. To evaluate the robustness against the second imposter attack scenario, we group the 100 subjects into 50 pairs. Each subject was told to mimic his/her partner's facial appearance and motion style and try to imitate him or her. One participant of the pair acted as an imposter, the other one as a genuine user. The genders of the imposter and the user were the same. They observed the facial appearance motion style of the target visually, which can be easily done in a real-life situation as facial appearance motion cannot be hidden. The authentication accuracy is evaluated by False Positive Rate (FPR) and False Negtive Rate (FNR). In general, FPR relates to the security of the system, while FNR to the usability. An interesting point in the Decision Error Trade-off (DET) curve is the Equal Error Rate (EER) where FPR = FNR. For instance, an EER of 1% means that out of 100 genuine trials 1 is incorrectly rejected, and out of 100 imposter trials 1 are wrongfully accepted. We vary the confidence threshold of the DNN predictions in *RFaceID* to plot DET curve in Fig. 23.

As shown in the figure, in both passive and active attacks, higher threshold makes the system more secure (i.e., with lower FPR), while also increases the FNR which requires users to shake their faces more times to get authenticated. The EER is 0.065 for passive attacks and is only 0.013 for active attacks. The EER of active attack is lower dues to the fact that in order to successfully launch an active attack, the imposter needs to mimic and get authenticated as one specific user, and this is harder than passive attacks. Overall, though the results indicate that attackers indeed have chances to get authenticated under passive and active attacks, the chance is still low. Besides, to lower FPR and improve the security, one can still increase the confidence threshold by trading off FNR based on application configurations. Other extensions are also possible to further enhance the security features as discussed in Section 5.

4.9 User Study

To evaluate the usability of *RFaceID*, we launch a questionnaire survey of 100 volunteers who used the *RFaceID* system. Fig.24 shows all the questions in our questionnaire and the corresponding survey results. The results show that 73% and 83% of volunteers think that the face recognition system like *RFaceID* is convenient and are willing to use it. 42% and 51% of volunteers think that *RFaceID* is easy to deploy and comfortable to use. Though *RFaceID* requires users to perform additional actions such as shaking faces comparing with conventional

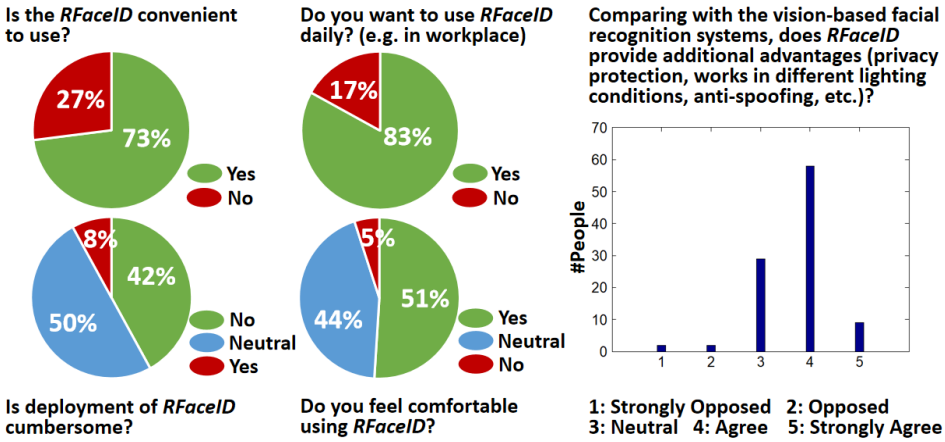


Fig. 24. Usability study on *RFaceID*

vision-based facial recognition system, 67% of the volunteers believe that *RFaceID* provides additional advantages such as privacy protection, little dependency on lighting conditions, and anti-spoofing, etc. The user study shows that *RFaceID* has sufficient usability and has potential to become one alternative approach to enhance current facial recognition systems.

5 LIMITATIONS AND FUTURE WORK

The *RFaceID* proposes a new RFID-based facial recognition approach. A RFID reader, antenna, and a RFID tag matrix are deployed to collect the RFID physical layer information for face recognition. Comparing with the state-of-the-art vision-based approach using a single camera, the deployment of *RFaceID* is more expensive and incurs additional hardware costs. Due to additional hardware dependency, *RFaceID* also cannot be easily integrated with the current devices such as smartphones or laptops, and the requirement for pre-deployment of RFID devices also restricts the application scenario of *RFaceID*. Comparing with the current vision-based approach, the face recognition accuracy of *RFaceID* is lower and is prone to face appearance changes such as wearing a mask as reported in Section 4. Besides, natural facial variations caused by a beard, moustache, etc. might also degrade the system performance and need to be explored in future work.

Despite the above limitations, *RFaceID* still has multiple benefits in terms of privacy protection, no dependency on lighting conditions, etc., which make it a novel complementary approach to enhance the current existing facial recognition systems. Multiple extensions are possible to integrates with *RFaceID* to further improve its usability in practice. (1) *Integrating with other sensor signals.* Other signals such as WiFi and mmWave radar can be integrated with *RFaceID* to further improve the accuracy, robustness, and security of the system, especially when users are wearing hats or face masks. (2) *Enhancing the security features.* Under the current threat model of *RFaceID* we assume the RFID signal collected is trustworthy, however under the threat model where attackers are able to inject adversarial examples and launch adversarial attacks, e.g., by injecting synthesized RSS/phase signals to the system, neural network protection techniques such as FGSM adversarial training [9] can be integrated to further enhance the security features of the system.

6 CONCLUSIONS

In this paper, we propose *RFaceID*, a novel RFID-based face recognition system. *RFaceID* is device-free and requires no privacy intrusive image/video input. By incorporating a set of novel data augmentation techniques and deep learning techniques, the *RFaceID* reduces the impact on the environmental noises and user dynamics during face recognition. *RFaceID* achieves a high recognition accuracy at 93.1%, and security analysis and user studies also show the usability of *RFaceID* in practice. Since *RFaceID* achieves accurate face recognition using RFID signals, it has the potential to open up a new range of future RFID-based facial recognition applications.

ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for their valuable comments. This work is supported by National Natural Science Foundation of China (61972263, U1713212, 62073225), Natural Science Foundation of Guangdong Province (2019A1515011608), and the Stable Support Plan for Higher Education Institutions in Shenzhen (20200810113310001).

REFERENCES

[1] Giuseppe Amato, Fabrizio Falchi, Claudio Gennaro, Fabio Valerio Massoli, Nikolaos Passalis, Anastasios Tefas, Alessandro Trivilini, and Claudio Vairo. 2019. Face Verification and Recognition for Digital Forensics and Information Security. In *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*. 1–6. https://doi.org/10.1109/ISDFS.2019.8757511

[2] Hind Baqeel and Saqib Saeed. 2019. Face detection authentication on Smartphones: End Users Usability Assessment Experiences. In *2019 International Conference on Computer and Information Sciences (ICCIS)*. 1–6. https://doi.org/10.1109/ICCISci.2019.8716452

[3] Ivana Chingovska, Nesli Erdogmus, André Anjos, and Sébastien Marcel. 2016. *Face Recognition Systems Under Spoofing Attacks*. Springer International Publishing, Cham, 165–194. https://doi.org/10.1007/978-3-319-28501-6_8

[4] Eduardo Costa, Ana Lorena, ACPLF Carvalho, and Alex Freitas. 2007. A review of performance evaluation measures for hierarchical classifiers. In *Evaluation methods for machine learning II: Papers from the AAAI-2007 workshop*. 1–6.

[5] Daniel Mark Dobkin. 2012. The RF in RFID: UHF RFID in Practice. (2012).

[6] Xiaoyi Fan, Wei Gong, and Jiangchuan Liu. 2018. TagFree Activity Identification with RFIDs. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 7 (March 2018), 23 pages. https://doi.org/10.1145/3191739

[7] X. Fan, F. Wang, W. Gong, L. Zhang, and J. Liu. 2018. Multiple Object Activity Identification Using RFIDs: A Multipath-Aware Deep Learning Solution. In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*. 545–555. https://doi.org/10.1109/ICDCS.2018.00060

[8] G. Forestier, F. Petitjean, H. A. Dau, G. I. Webb, and E. Keogh. 2017. Generating Synthetic Time Series to Augment Sparse Datasets. In *2017 IEEE International Conference on Data Mining (ICDM)*. 865–870. https://doi.org/10.1109/ICDM.2017.106

[9] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. 2014. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572* (2014).

[10] J. Han, C. Qian, X. Wang, D. Ma, J. Zhao, P. Zhang, W. Xi, and Z. Jiang. 2014. Twins: Device-free object tracking using passive tags. In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*. 469–476. https://doi.org/10.1109/INFOCOM.2014.6847970

[11] Impinj. 2010. R420 readers. (2010). Retrieved June 16,2020. https://www.impinj.com/library.

[12] Sergey Ioffe and Christian Szegedy. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *Proceedings of the 32nd International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 37)*, Francis Bach and David Blei (Eds.). PMLR, Lille, France, 448–456. http://proceedings.mlr.press/v37/ioffe15.html

[13] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[14] A. Krizhevsky, I. Sutskever, and G. Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In *NIPS*.

[15] Xinyu Li, Yanyi Zhang, Ivan Marsic, Aleksandra Sarcevic, and Randall S. Burd. 2016. Deep Learning for RFID-Based Activity Recognition. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM* (Stanford, CA, USA) *(SenSys '16)*. Association for Computing Machinery, New York, NY, USA, 164–175. https://doi.org/10.1145/2994551.2994569

[16] Jian Liu, Hongbo Liu, Yingying Chen, Yan Wang, and Chen Wang. 2020. Wireless Sensing for Human Activity: A Survey. *IEEE Communications Surveys Tutorials* 22, 3 (2020), 1629–1645. https://doi.org/10.1109/COMST.2019.2934489

[17] Zhuo Ma, Yang Liu, Ximeng Liu, Jianfeng Ma, and Kui Ren. 2019. Lightweight privacy-preserving ensemble classification for face recognition. *IEEE Internet of Things Journal* 6, 3 (2019), 5778–5790.

[18] M. Schuster and K. K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* 45, 11 (1997), 2673–2681. https://doi.org/10.1109/78.650093

[19] Daniel Sáez Trigueros, Li Meng, and Margaret Hartnett. 2018. Face Recognition: From Traditional to Deep Learning Methods. arXiv:1811.00116 [cs.CV]
[20] Leung Tsang, Jin Au Kong, and Robert T Shin. 1985. Theory of microwave remote sensing. (1985).
[21] Weiye Xu, Jianwei Liu, Shimin Zhang, Yuanqing Zheng, Feng Lin, Jinsong Han, Fu Xiao, and Kui Ren. 2021. RFace: Anti-Spoofing Facial Authentication Using COST RFID. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*.
[22] Lei Yang, Qiongzheng Lin, Xiangyang Li, Tianci Liu, and Yunhao Liu. 2015. See Through Walls with COTS RFID System!. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking* (Paris, France) *(MobiCom '15)*. Association for Computing Machinery, New York, NY, USA, 487–499. https://doi.org/10.1145/2789168.2790100
[23] Lin-Lin Zhang, Jing Xu, Daim Jung, Tabe Ekouka, and Ha-Kyun Kim. 2021. The Effects of Facial Recognition Payment Systems on Intention to Use in China. *Journal of Advanced Researches and Reports* 1, 1 (2021), 33–40.
[24] C. Zhao, Z. Li, T. Liu, H. Ding, J. Han, W. Xi, and R. Gui. 2019. RF-Mehndi: A Fingertip Profiled RF Identifier. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*. 1513–1521. https://doi.org/10.1109/INFOCOM.2019.8737419