Anti-Spoofing Facial Authentication Based on COTS RFID

Weiye Xu, *Student Member, IEEE*, Jianwei Liu, *Student Member, IEEE*, Shimin Zhang, Yuanqing Zheng, *Member, IEEE*, Feng Lin, *Senior Member, IEEE*, Fu Xiao, *Member, IEEE*, and Jinsong Han, *Senior Member, IEEE*

Abstract—Current facial authentication (FA) systems are mostly based on the images of human faces, thus suffering from privacy leakage and spoofing attacks. Mainstream systems utilize facial geometry features for spoofing mitigation, but they are still vulnerable to feature manipulation, *e.g.*, 3D-printed human faces. In this paper, we propose a novel privacy-preserving anti-spoofing FA system, named RFace, which extracts both the 3D geometry and inner biomaterial features of faces using a COTS RFID tag array. These features are difficult to obtain and forge, hence are resistant to spoofing attacks. Unlike images, RF signals are not perceptible to human eyes, so RFace protects user's privacy. We build a theoretical model to rigorously prove the feasibility of feature acquisition and the correlation between facial features and RF signals. To enhance the security of RFace, we specify the tag reading order for each authentication to defend against the signal replay attack. For practicality, we design an effective algorithm to mitigate the impact of unstable distance and angle deflection from the face to the array. Extensive experiments with 30 participants and three types of spoofing attacks show that RFace achieves an average authentication success rate of over 95.7% and an EER of 4.4%. More importantly, no replay attack or spoofing attack succeeds in deceiving RFace in the experiments.

Index Terms—RFID, Wireless Sensing, Internet of Things, Facial Authentication.

1 Introduction

In recent years, facial authentication (FA) systems have been wildly applied to daily applications (*e.g.*, access control, online payment and individual identification [1, 2, 3]). FA is regarded as a promising alternative to traditional authentication approaches, such as PIN code [4], fingerprint [5] and token [6] thanks to its convenience and precision.

Existing FA systems are mostly camera-based and have some severe flaws, including privacy leakage risks and security problems. On the one hand, most current FA systems collect facial features of users by RGB cameras, which inevitably reveals complete visual facial information (VFI) and raises privacy concerns [7, 8]. Although some methods have been proposed to preserve users' privacy, e.g., applying the cryptographic protocol [9, 10, 11] and adding perturbation [12], they can not work without a trusted third party, thus can not solve the privacy issues fundamentally. On the other hand, these approaches perform authentications by extracting geometry features from VFI. Once the VFI of a user is leaked, an attacker can easily conduct a spoofing attack by reproducing the features [13]. Although more geometry information has been introduced to enhance the security, *e.g.*, the depth information of faces [14], camera-based FA systems are still vulnerable to spoofing attacks. This is because the information can still be captured remotely, *e.g.*, using depth cameras or infrared dot projectors [15]. Therefore, attackers can manipulate a 3D-printed mask based on the captured VFI to deceive FA systems [16]. To overcome these drawbacks, we explore a new facial feature based authentication technique, which can protect the visual privacy of users, and meanwhile resist spoofing attacks.

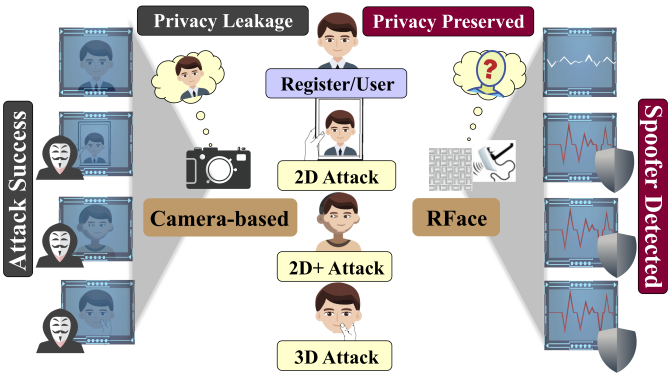


Fig. 1: Compared with camera-based systems, *RFace* is privacy-preserving and can effectively resist spoofing attacks.

We turn our attention to radio frequency (RF) sensing. Recent advances in wireless sensing reveal that radio frequency (RF) signals are sensitive to the material that they encounter during propagation [17**?**]. Thus, by verifying the facial biomaterial feature captured in RF signals, we can determine the authenticity of users and defend against spoofing attacks. Meanwhile, unlike images that leak the user's VFI, RF signals are not perceptible to the human eye and can effectively protect the users' privacy. Moreover,

Weiye Xu, Jianwei Liu, Shimin Zhang are with Zhejiang University, Hangzhou, China, 310027, and also with the ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou, China, 311215. Email: xuweiye@zju.edu.cn, jianweiliu@zju.edu.cn, shiminzsm@gmail.com. Feng Lin and Jinsong Han are with the Zhejiang University, Hangzhou, China, 310027, and also with the Key Laboratory of Blockchain and Cyberspace Governance of Zhejiang Province, China, 310027. Email: flin@zju.edu.cn, hanjinsong@zju.edu.cn.
Yuanqing Zheng is with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: yqzheng@polyu.edu.hk. Fu Xiao is with the College of Computer, Nanjing University of Posts and Telecommunications, China, 210003. Email: xiaof@njupt.edu.cn.
Jinsong Han is the corresponding author.

researchers show that an array of commercial-off-the-shelf (COTS) radio frequency identification (RFID) tags can serve as a low-cost and sensitive sensor to support fine-grained wireless sensing [18, 19]. Motivated by the advantages of RFID, we attempt to develop an anti-spoofing and privacy-preserving FA system with COTS RFID devices.

In this paper, we propose a novel anti-spoofing privacy-preserving authentication system, named RFace. Specifically, RFace uses an RFID tag array to measure both RSS and phase of RF signals and extract 3D facial geometry and biomaterial features for spoofing attack resistance. The extracted features are then fed into a well-trained support vector machine (SVM) to conduct authentication. As shown in Fig. 1, compared with camera-based systems, RFace protects the VFI of users. RFace can effectively resist spoofing attacks since the extracted features are difficult to be captured by attackers.

There are three main challenges in the design and development of RFace: (1) It is difficult to build the correlation model between real facial features and raw RF signals since RF signals are not as structured as images. (2) RF-based authentication system is prone to signal replay attacks. Previous works [20] show that if an attacker can replay the exact similar RF signals of a legitimate object (e.g. a human face), the authentication method without corresponding defense mechanisms can be deceived. (3) For each user, the distance and deflection from the face to the tag array may slightly vary in different authentication attempts, leading to undesirable degradation of the authentication accuracy due to the sensitivity of RF signals.

To address the first challenge, we build a theoretical model based on the principles of electromagnetism and RF signal propagation. The model rigorously proves the feasibility of leveraging RF signals to capture and extract the desired facial features from raw backscatter signals. To tackle the second challenge, we design a replay attack defense method that specifies the reading order of tags for each authentication to increase the difficulty of conducting replay attacks. While enhancing the security of the FA system, we also take the efficiency of collecting RF signals into consideration by designing an efficient reading scheme. For the third challenge, we try to explore a feature resistant to the variation of distance and deflection angle. Based on experiments and theoretical analysis, we find that in the tag array, the difference values of RSS and phase of two tags are related to the distance between the two tags. Specifically, the RSS and phase differences of two tags that are close to each other are more stable than that of two tags with a far distance, especially when the distance and deflection angle vary. Therefore, we take the RSS and phase differences of two tags within a tiny area as the final stable feature. We propose a novel algorithm, named DR, to split the array into the smallest blocks, *i.e.*, a tag with its adjacent tags. In this way, the adjacent tags leverage their stable difference values of RF signals towards the variation of distance and deflection, to mitigate the degradation in the authentication accuracy.

RFace is implemented with COTS RFID devices. To perform user authentication, users are only required to pose their faces in front of a tag array for a few seconds. We evaluate RFace with 30 volunteers under various distance and deflection conditions. We also test RFace against replay attack and three types of mainstream spoofing attacks, including 2D, 2D+ and 3D spoofing attacks. The results demonstrate the effectiveness of RFace with over 95.7% authentication success rates and around 4.4% equal error rates. More importantly, the results of defense experiments indicate that RFace is resistant to replay attack and spoofing attacks, including 3D mask attacks. Therefore, we can envision that RFace has the potential to complement existing FA systems in some challenge scenarios, e.g., protecting privacy and defending against spoofing attacks. Our contributions can be summarized as follow:

- We propose a novel privacy-preserving anti-spoofing FA system, RFace, which can extract both 3D facial geometry and biomaterial features of users' faces from RF signals. In addition, we build a theoretical model to validate the feasibility of the feature extraction method.
- We propose a replay attack defense approach to enhance the security of RFace while ensuring the efficiency of our system.
- We propose a novel algorithm to enhance the flexibility of RFace by mitigating the impact of distance and deflection variations between the human face and tag array.
-

2 PRELIMINARY

In this section, we start with some preliminary studies on two elemental indicators of RF signals and the layout of the tag array that facilitates the feature collection. Then, we introduce four types of typical attacks on FA systems.

2.1 Indicators of RF Signals and Tag Array Layout

RF Signal Indicators. A typical RFID system includes three parts: RFID tags, a reader and its antenna. The reader continuously transmits RF signals to activate tags. Each tag responds its Electronic Product Code (EPC) by backscattering RF signals using ON-OFF keying [21]. By analyzing the backscatter signals, the reader can obtain the indicators, *i.e.*, received signal strength (RSS) and phase. Both indicators are influenced by the propagation distance and the materials encountered. In this paper, RFace measures the indicators to extract facial features. The concrete mathematical model will be presented in Section 3.1.

In this work, we utilize the *Impinj Speedway R420* reader which can achieve a phase resolution of 0.0015 radians in theory [22]. Thus, with the corresponding wavelength about $32cm$, it is capable of providing $\approx 320mm * 0.0015/(4 \times 3.14) = 0.038mm$ distance resolution, which suffices to capture the geometry feature of human face.

Tag Array Layout. To alleviate the inductive coupling effect, we adopt a perpendicular orientation deployment of tags as in [19]. Additionally, for avoiding two tags on a tag array from sharing the same phase, we ensure the maximum distance between any two tags on an array is less than half of the wavelength. Specifically, in RFace, we use the tiny RFID tags *AZ-9629* (only $2.25cm \times 2.25cm$) to arrange a compact tag array with the geometry of square space, which is capable to cover a whole face.

2.2 Attacks

We mainly consider the following four types of attacks against FA systems.

2D Spoofing Attack: This attack includes both 2D static (photo) attack and dynamic (video) attack. 2D spoofing attacks aim to deceive an FA system with a 2D photo/video of a legitimate user's face. Most traditional 2D feature based FA systems are vulnerable to this attack. As reported in [23], a 2D photo can deceive the face recognition systems of some mainstream smartphones.

2D+ Spoofing Attack: This attack constructs an uneven mask based on a precise 2D image plus rough depth measurements of a legitimate user's face. Such an attack can deceive many existing anti-spoofing FA systems [2] that only look for uneven surfaces rather than matching against the facial depth information exactly.

3D Spoofing Attack: This attack constructs a precise 3D-printed mask of a legitimate user. 3D spoofing attacks are hard to defeat, since current FA systems only extract the 3D structural features of the face surface, while ignoring the inner composition, *e.g.*, the biomaterial of faces. In this case, using a vivid 3D-printed mask can easily trick these FA systems [24].

Replay Attack: This attack eavesdrops on the RF signals of legitimate objects (e.g., a legitimate user's face), and then replays the captured signal to deceive the FA system. One common method to intercept signals is to use an unauthorized reader to capture the backscatter signals of tags during the registration/authentication period. This kind of attack is particularly aimed at wireless authentication systems that utilize RF signals for verification [20].

3 Turning RFID Into Facial Features Extractor

In order to establish the correlation between RF signals and both the 3D geometry and inner biomaterial features of faces, we build a theoretical model based on the principles of electromagnetism and RF signal propagation. Besides, we conduct several experiments to test the feasibility of anti-spoofing face authentication.

3.1 Modeling Face Features upon RF Signal Propagation

Modeling Consideration: In our system, the received signals consist of three parts: the line-of-sight signal, the signal reflected from other surrounding objects and the signal reflected from face. The line-of-sight signal can be regarded as a consistent one when the distance between the antenna and the tag array remains static. In the FA scenario, the distance between other surrounding objects and the tag array is normally much larger than the one between the face and the tag array. As a result, the face would contribute the most of the impact on the tag array in terms of RF interaction. Therefore, we focus on the signal reflected from the face in our model. Moreover, we utilize the RSS and phase of received signals to extract the user's facial features. We demonstrate that the extracted features can represent both the unique *3D geometry* and *inner biomaterial* features of human face in the following.

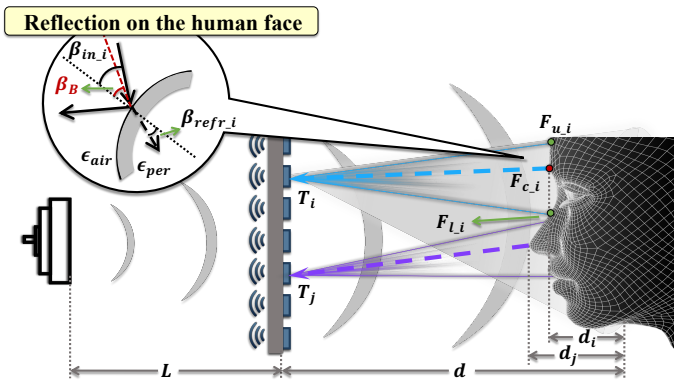


Fig. 2: RF signal propagation model.

As illustrated in Fig. 2, the human face is located directly in front of and parallel to the tag array. The user's chin is located on the vertical bisector of the bottom edge of the tag array. We utilize the lateral view of a 7×7 tag array to interpret the effect of the signal reflected from the face on the tag array. We denote the distance between the antenna and the tag array as L and the distance between the tag array and the chin as d. Our theoretical model involves three processes: 1) mapping tags to blocks on face, 2) extracting facial features by RSS, and 3) extracting facial features by phase.

Mapping Tags to Blocks on Face. In order to precisely extract facial features, it is essential to map tags to their corresponding areas on a face. Due to the sensitivity of RF signal, each tag on the tag array backscatters the reflection signal from the entire face. Thus, each tag captures the features of the entire face. Nevertheless, the face will have a greater impact on those tags closer to the face [19]. Therefore, without losing the effectiveness of modeling, for each tag, we analyze the influence on the tag from the area on the face which is closest to the tag. In specific, we map each tag on the array to an area closest to the tag on the face, which can be regarded as a block. For ease of modeling, this reflection can be equivalent to the effect of the center point of this block on the tag. Taking tag T_i as an example, the corresponding reflection block is represented as the closest area (F_{u_i}, F_{c_i}, F_{l_i}), where F_{u_i}, F_{c_i}, F_{l_i} represent the upper vertex, center point and lower vertex of this block on the face, respectively. Therefore, the signal reflected from this block and received by tag T_i can be regarded as the signal reflected from the center point F_{c_i} and we denote the incident angle of RF signal propagating into human face at F_{c_i} as β_{in_i}. In addition, for distinct human faces, each tag on tag array has different specific reflection area with corresponding center point and incident angle. Then we will take tag T_i and tag T_j as an instance to analyze the impact of RF signals reflected from human face on RSS and phase, and show how to extract facial features.

Facial Features Extracted by RSS. We divide the propagation of RF signals in our system into three stages: from antenna to face, reflection on face and from face to antenna. We analyze the change of RSS in each stage respectively to extract both 3D geometry and inner biomaterial features.

As for the first propagation stage, the RSS value is defined by the power P, which is proportional to the square

of the amplitude A. The RSS can be formulated as:

$$RSS = 10\log\frac{P}{10^{-3}} = 20\log(DA) \quad (P \propto A^2), \tag{1}$$

where D is a constant. Generally, the amplitude A causes exponential loss in magnitude over a unit of propagation distance, which can be denoted as $e^{-\alpha}$. For the tag T_i in Fig. 2, the loss of amplitude A during the propagation of RF signal from antenna to the center point F_{c_i} of the corresponding reflection block (F_{u_i}, F_{c_i}, F_{l_i}) on face can be denoted as:

$$A_{in} = Ae^{-\alpha(L+d-d_i)}. \tag{2}$$

In this equation, A_{in} represents the amplitude of the RF signal when it reaches the surface of face. d_i denotes the horizontal distance between the center point F_{c_i} and chin, which reveals the 3D facial geometry features of human face.

For the second propagation stage, the RF signal reaches the surface of human face. The signal can be divided into two parts: the part directly reflected and the part entering the face with refraction. The inner structure of human face consists of various biomaterials such as skin, fat, and muscle. For ease of illustration, we assume that each person's face is a unique hybrid material composed of multiple layers of materials, and the relative permittivity of the mixed biomaterial can be denoted as ϵ_{per}. According to [17], the RF signal refracted into the face has to traverse multi-layered tissue and go through multiple reflections before it can escape. Due to the exponential attenuation, there will be only very low power signal to escape the human face, which can be ignored. Then, we mainly analyze the power loss caused by reflection on the face surface. The power ratio of the signal before and after reflection on the face surface is the power reflection coefficient R_{per_i}, so the corresponding amplitude can be calculated as:

$$A_{\_after} = \sqrt{R_{per\_i}}A_{\_before}, \tag{3}$$

where $A_{\_before}$ and $A_{\_after}$ are the amplitudes of the corresponding before and after reflection signal. R_{per_i} is related to the mixed material of human face and the incident angle β_{in_i}. So R_{per_i} can reveal the facial features of human face. Then, we analyze the factors that affect R_{per_i} in detail.

The reflection on an interface between two biomaterials is affected by the relative permittivity ϵ of them. Considering the RF signal arriving at the face surface which is the interface between the air (ϵ_{air}) and the face material (ϵ_{per}), the signal can be decomposed into transverse electric (TE) and transverse magnetic (TM) wave components, according to [25]. The power reflection coefficient R_{per_i} can be calculated by R_s and R_p, which are the power reflection coefficient of TE and TM wave components respectively. According to Fresnel formula [25], R_p and R_s can be calculated as:

$$\begin{cases} TE: R_s = \left|\frac{\sqrt{\epsilon_{per}}cos\beta_{refr\_i} - \sqrt{\epsilon_{air}}cos\beta_{in\_i}}{\sqrt{\epsilon_{per}}cos\beta_{refr\_i} + \sqrt{\epsilon_{air}}cos\beta_{in\_i}}\right|^2 \\ TM: R_p = \left|\frac{\sqrt{\epsilon_{per}}cos\beta_{in\_i} - \sqrt{\epsilon_{air}}cos\beta_{refr\_i}}{\sqrt{\epsilon_{per}}cos\beta_{in\_i} + \sqrt{\epsilon_{air}}cos\beta_{refr\_i}}\right|^2 \end{cases}, \tag{4}$$

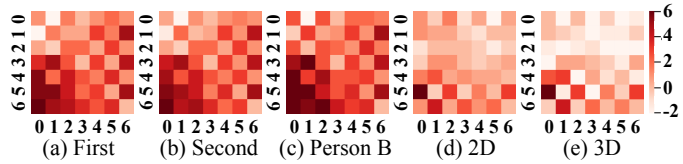


Fig. 3: RSS difference distributions.

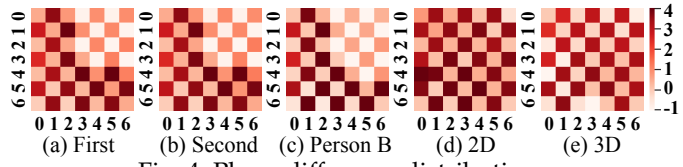


Fig. 4: Phase difference distributions.

where β_{refr_i} denotes the refraction angle and ϵ_{air} denotes the relative permmitivity of the air. R_{per_i} can be represented as:

$$\begin{aligned} R_{per\_i} &= R_s R_p \\ &= \left|\frac{\sqrt{\epsilon_{per}}cos\beta_{refr\_i} - \sqrt{\epsilon_{air}}cos\beta_{in\_i}}{\sqrt{\epsilon_{per}}cos\beta_{refr\_i} + \sqrt{\epsilon_{air}}cos\beta_{in\_i}}\right|^2 \\ &\times \left|\frac{\sqrt{\epsilon_{per}}cos\beta_{in\_i} - \sqrt{\epsilon_{air}}cos\beta_{refr\_i}}{\sqrt{\epsilon_{per}}cos\beta_{in\_i} + \sqrt{\epsilon_{air}}cos\beta_{refr\_i}}\right|^2. \end{aligned} \tag{5}$$

Furthermore, according to Snell's Law [25], we have:

$$\sqrt{\epsilon_{per}}sin\beta_{refr\_i} = \sqrt{\epsilon_{air}}sin\beta_{in\_i}, \tag{6}$$

which means that β_{refr_i} can be represented by β_{in_i}, ϵ_{air} and ϵ_{per}. Therefore, we can conclude that R_{per_i} is only determined by ϵ_{per} and β_{in_i}. Therefore, in the propagation stage of reflection on face, the value of R_{per_i} involves both the inner biomaterial and 3D geometry features of human face.

Then for the last propagation stage from face to antenna, combining with Eq. 2 and Eq. 3, the amplitude of received signal A_r can be calculated as:

$$A_r = \sqrt{R_{per\_i}}Ae^{-2\alpha(L+d-d_i)}. \tag{7}$$

According to Eq. 1, the RSS of the received signal RSS_i for tag T_i can be denoted as:

$$RSS_i = 20\log(DA_r) = 20\log\left[D\sqrt{R_{per\_i}}Ae^{-2\alpha(L+d-d_i)}\right]. \tag{8}$$

We can find that the value of RSS is relevant to the distance (d) between the tag array and chin, and d is an unstable factor because of the movement of human face. In order to better extract stable facial features, we then subtract the RSS values of these two tags T_i and T_j to remove the impact of d, which is irrelevant to the structure of human face. Then we can gain:

$$\begin{aligned} RSS_i - RSS_j &= 20\log\frac{D\sqrt{R_{per\_i}}Ae^{-2\alpha(L+d-d_i)}}{D\sqrt{R_{per\_j}}Ae^{-2\alpha(L+d-d_j)}} \\ &= 20\log\left[\frac{\sqrt{R_{per\_i}}}{\sqrt{R_{per\_j}}}e^{-2\alpha(d_j-d_i)}\right]. \end{aligned} \tag{9}$$

It can be observed that the factor d has been eliminated and the value of $RSS_i - RSS_j$ only related to R_{per_i}, R_{per_j} and $d_j - d_i$. Among these variables, R_{per_i} and R_{per_j} involve both the inner biomaterial feature and 3D geometry feature of human face which vary among persons. Meanwhile

$d_j - d_i$ represents the horizontal distance between the center points of two different blocks on the human face, which also indicates the 3D geometry feature of human face. Therefore, we can conclude that the difference between the RSS values of the tags can reveal both distance-resistant 3D geometry and inner biomaterial features of human face.

Facial Features Extracted by Phase: Similar to RSS, we analyze the changes of phase according to the three stages of RF signal propagation in our model successively. Taking tag T_i as an example, the phase of the RF signal transmitted from the antenna before entering face ($\theta_{\_before}$) can be expressed as:

$$\theta_{\_before} = \theta_t + \frac{2\pi}{\lambda}(L + d - d_i), \tag{10}$$

where θ_t represents the initial phase value of the transmitter. As aforementioned, the signal is thought to be unable to escape once it enters the face, so only the reflections that occur on the surface of human face are considered. In this case, the phase of the RF signal ($\theta_{\_after}$) after reflection can be calculated as:

$$\theta_{\_after} = \theta_{\_before} + \theta_{per\_i}, \tag{11}$$

where θ_{per_i} denotes the phase variation caused by reflection at the human face. We then show that θ_{per_i} is also related to the facial features. According to [25], after reflections, there will be a phase change when the angle of incidence (β_{per_i}) is larger than Brewster angle (β_B):

$$\begin{cases} \theta_{per\_i} = 0, & \beta_{per\_i} \le \beta_B \\ \theta_{per\_i} = \pi, & \beta_{per\_i} > \beta_B \end{cases}, \beta_B = arctan\sqrt{\frac{\epsilon_{per\_i}}{\epsilon_{air}}}. \tag{12}$$

Therefore, θ_{per_i} is decided by ϵ_{per_i} and β_{per_i}, which means that the value of the phase variation (θ_{per_i}) caused by reflection on human face can reveal both the inner biomaterial and 3D facial geometry features of the face.

Combining with the phase changes in the third stage and substituting Eq. 10 into Eq. 11, we can obtain the final phase of the received RF signal for tag T_i as:

$$\begin{aligned} \theta_i &= \left[\theta_{\_after} + \theta_{tag\_i} + \frac{2\pi}{\lambda}(L + d - d_i)\right] \mod 2\pi \\ &= \left[\theta_t + \theta_{per\_i} + \theta_{tag\_i} + \frac{4\pi}{\lambda}(L + d - d_i)\right] \mod 2\pi, \end{aligned} \tag{13}$$

where θ_{tag_i} denotes the additional phase shift caused by the characteristic of tag. Similarly, to remove the factor d, we subtract θ_i with θ_j and the result can be denoted as:

$$\begin{aligned} \theta_i - \theta_j =& (\theta_{per\_i} - \theta_{per\_j}) + (\theta_{tag\_i} - \theta_{tag\_j}) \\ &+ \frac{4\pi}{\lambda}(d_j - d_i) + 2k\pi, k \in Z. \end{aligned} \tag{14}$$

It is worth noting that there is a term $2k\pi$ in theory since the phase of the received RF signal is a periodic function with a period of 2π. This uncertainty term can be eliminated when we apply a cosine function in Section 5. Similar to the analysis of RSS, we can conclude that the difference value of phase between tags ($\theta_i - \theta_j$) can represent the 3D geometry and inner biomaterial features of human faces.

3.2 Feasibility of Anti-Spoofing Face Authentication

We conduct experiments to demonstrate that the 3D geometry and inner biomaterial features of human face have been indeed extracted for anti-spoofing FA purpose. Specifically, we ask two volunteers (Person A and Person B) to put their faces at a fixed distance in front of the tag array twice, and record the phase and RSS values of the tags. Then, we hold the 2D photo and 3D printed mask of the two volunteers respectively, and collect the phase and RSS values. In Fig. 3 and Fig. 4, we calculate the phase difference and the RSS difference between the surrounding tags with the center one to analyze the impact of different targets on the tags. We can observe that the result of 2D photo is different from that of 3D mask, (Fig. 3(d) vs. Fig. 3(e) and Fig. 4(d) vs. Fig. 4(e)), which shows that the 3D geometry features of human face have been extracted. Besides, the result of 2D photo also varies severely to the real human face (Fig. 3(a) vs. Fig. 3(d) and Fig. 4(a) vs. Fig. 4(d)), which indicates that the traditional 2D static and dynamic attacks can be defended. Similarly, the result of 3D mask is distinct to the real face (Fig. 3(a) vs. Fig. 3(e) and Fig. 4(a) vs. Fig. 4(e)), indicating that the inner biomaterial features of human face have been extracted, and our system can resist 3D spoofing attacks. Additionally, the results of the same volunteer at different attempts are very similar (Fig. 3(a) vs. Fig. 3(b) and Fig. 4(a) vs. Fig. 4(b)), while they are different for two volunteers (Fig. 3(a) vs. Fig. 3(c) and Fig. 4(a) vs. Fig. 4(c)). Therefore, the features we extract are unique to everyone, so they can be utilized for authentication. Additionally, since these characteristics of a specific human face are difficult to imitate, our system can resist spoofing attacks.

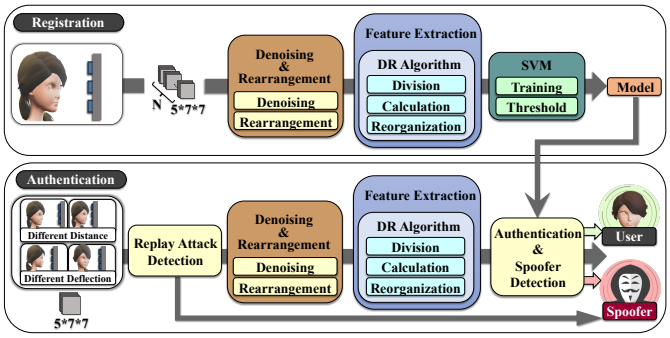


Fig. 5: System architecture of *RFace*.

4 System Overview

This section first shows the overview of RFace system and then illustrates its usage scenario.

4.1 Overview

The architecture of RFace consists of two primary phases (in Fig. 5): the registration phase and the authentication phase. Users can register and authenticate by simply posing their faces in front of a tag array for a few seconds. In the registration phase, RFace collects the RSS and phase values of the RF signals reflected from human face, and feeds them into a denoising and rearrangement algorithm to construct an RSS and phase sequence. Then, this sequence goes through a disturbance resistant facial features extraction algorithm. The reliable fusion features consisting of 3D geometry and inner biomaterial can be extracted by calculating the RSS and phase differences between the tags on the tag array. Finally, the extracted fusion features are organized into

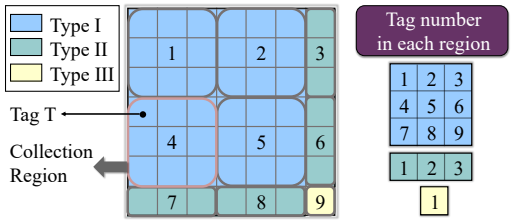


Fig. 6: Dividing the tag array into regions.

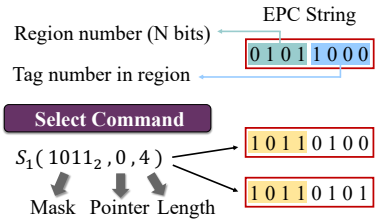


Fig. 7: EPC string and select command.

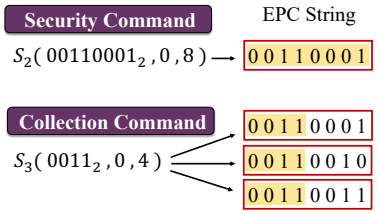


Fig. 8: Security command and collection command.

feature blocks according to the time dimension and fed into an SVM to train a user classifier.

In the authentication phase, the collected RF signals must first go through the replay attack detection module. Once a real human face is detected, RFace randomly generates an EPC sequence to specify the reading order of tags and transmits the corresponding reading commands to query tags. According to the reading commands, RFace starts to collect the RSS and phase values of the backscatter RF signals and compares the actual reading order of tags with the preset EPC sequence. If they can not match, we will regard this authentication as a replay attack and directly reject this request. If they match successfully, the collected RF signals will further go through a denoising and facial feature extraction process. Afterwards, RFace leverages the extracted fusion features to conduct authentication and spoofing attack defense.

4.2 Usage scenario

RFace is mainly designed for providing FA in public institutions, such as access control of group clients (e.g., communities, small- or moderate-scale companies) and individual identification of smart home devices [26]. Considering the salient features of RF sensing including human-eye imperceptibility and material sensitivity, RFace can also be applied to complement and enhance existing camera-based methods in some challenging scenarios, e.g., suffering from privacy leakage and spoofing attacks. Furthermore, using RFace for authentication, users only need to put their faces in front of the tag array for a short time, which is as convenient as existing camera-based methods.

5 SYSTEM DESIGN

In this section, we present the design details of RFace. Before authentication, we first perform replay attack detection based on the collected RF signals. Then, we remove noise from raw RF signals and suppress the impacts of distance and deflection variations between the face and the tag array to extract stable fusion features including both 3D geometry and inner biomaterial features. Finally, we utilize an SVM to realize anti-spoofing authentication.

5.1 Replay Attack Defense

Replay attack, as one of the most tricky attacks for wireless authentication systems, threatens the security of RFace. Existing works to defend against replay attacks typically add random noise [20], which will destroy the features embedded in RF signals and affect the accuracy of authentication. Therefore, we attempt to develop an approach that can resist replay attacks while preserving high-quality facial features. To this end, we randomly generate an EPC sequence to specify the reading order of tags before each authentication, and then compare the actual reading order of tags with the preset EPC sequence to detect replay attacks. When a legitimate user utilizes RFace for authentication, the tags will be read sequentially according to the preset EPC sequence. In this case, the actual reading order has a high similarity with the preset EPC sequence. On the contrary, an attacker may try to deceive our system by eavesdropping the RF signals of a legitimate user while using RFace and replaying the RF signals later. Since the EPC sequence is generated randomly for each authentication, the probability that the actual reading order of the replayed RF signals can exactly match the current preset EPC sequence is extremely low. Transforming this high-level idea into our practical FA system requires tackling three main challenges. (1). Balancing security and tag reading efficiency. To increase the difficulty of implementing replay attacks, it is necessary to generate as many different EPC sequences as possible. The method to generate most EPC sequences is to read one tag each time, so for an $n * n$ tag array, there are $(n * n)!$ possible EPC sequences. Although this method can enhance the security of our system, it will significantly impact the efficiency, since it requires sending n reading commands to read all tags on the tag array. (2). Ensuring the stability of local features. During practical authentication, a user's face may sway unconsciously. If the reading time interval between adjacent tags on a tag array is too long, the local facial features captured by RFace would be destroyed, which will affect the accuracy of authentication. (3). Resilience to bit errors. Due to the imperfection of hardware, it is possible to miss reading the backscatter signals of tags. Therefore, even for legitimate users, the actual reading order of tags may not match the preset EPC sequence. This situation is called bit error. Bit error may make RFace mistakenly regard a legitimate user as an attacker. To address the above challenges, we design an **efficient reading scheme** to read all tags on the tag array with fewer commands while ensuring security. In order to make this scheme better applicable to real scenarios, we propose a **replay attack detection** method that is resilient to bit errors.

5.1.1 *Efficient Reading Scheme*

To improve the efficiency of the FA system, we hope to read a group of tags by sending one reading command, rather than only reading one tag per command. However, if the tags of the group (called *reading group*) read by one reading command are randomly arranged in the tag array, the local

facial features captured by the tag array will be affected. Therefore, we need to design a method that can make the tags in the *reading group* adjacent to each other spatially in the tag array. To this end, we propose an **efficient reading scheme**, which describes how to efficiently read tags while ensuring the security and stability of local facial features. Specifically, we first divide the tag array into several regions according to the spatial position. Then, as shown in Fig. 6, we randomly select a tag (Tag T) from all unread tags (initially all tags are unread) in the tag array as the reading target of the first reading command. For the subsequent command, the reading targets will be set as all tags on the region where Tag T locates. Afterwards, we repeat the above operations for the remaining unread tags until all tags have been read. We call the reading command whose reading target is only one tag as a *security command*. It specifies the reading order of tags to ensure the security of the FA system. We call the reading command whose reading targets are all tags on a region as a *collection command*. It is responsible for collecting the backscatter signals of a group of tags without caring about the reading order, thus ensuring efficiency. In RFace, these two commands are transmitted alternatively. We describe the details of this **efficient reading scheme** below.

Dividing tag array into regions. As aforementioned, to ensure the stability of local facial features, we need to divide the tag array into regions. In our system, we define three types of region templates and utilize them to divide the tag array. Specifically, as shown in Fig. 6, Type I represents a $3*3$ square region. Type II represents a region whose number of rows or columns is 3. Type III represents a region whose row number and column number are both less than 3. As shown in Fig. 6, taking a $7 * 7$ tag array as an example, we first utilize the Type I region template to divide it. The remaining tags that can not be divided by the Type I region template will be further divided by the Type II region template into new regions. After the above dividing process, the area of undivided tags whose row number and column number are both less than 3 will construct a new region (Type III). Then, as shown in Fig. 6, we sort these regions from top to bottom and from left to right, denoting them as 1, 2, 3 and so on. For tags in each region, we also sort them from top to bottom and from left to right, and number them as 1, 2, $3 \cdots 9$.

Security and collection commands. As aforementioned, one region contains nine tags at most so a four bits binary string (called tag number in region) can represent all tags in a region (as shown in Fig. 7). According to the number of regions, we can know how many bits are needed to represent all regions. For RFace, the $7 * 7$ tag array can be divided into nine regions, so we can utilize four bits (called region number) to represent all regions (as shown in Fig. 7). In this way, we can utilize an eight bits binary string to represent the EPC strings of each tag on a tag array.

The core of security and collection commands is to selectively read the tags of interest. In RFace, we realize these two commands by leveraging a universal Gen2 command '*Select*', which supports selecting target tags to be read [27]. The '*Select*' command consists of six required fields and one optional field. RFace focuses on three fields: *Mask*, *Pointer*, and *Length*, which are responsible for selecting target tags. The *Mask* field specifies a binary string for comparison. The *Pointer* field denotes the starting address on the EPC string for comparison. The *Length* field represents the length of the *Mask*. As shown Fig. 7, taking one *Select* command as an example, if *Mask* $= 1011_2$, *Pointer* $= 0$ and *Length* $= 4$, we will select the tags whose bit string starting from 1^{st} bit and ending at $(1 + 4)^{th}$ bit of the EPC string equals to 1011_2.

Then, we will show how to utilize the '*Select*' command to realize the security and collection commands. As for the *security command*, its reading target is only one tag, e.g. the 1^{st} tag of the 3^{rd} region, so the *Mask* needs to be set as the EPC string of the reading target. Therefore, as shown in Fig. 8, we set the *Length* $= 8$, *Pointer* $= 0$, and *Mask* to the EPC string of target tag, i.e., *Mask* $= 00110001_2$. For the collection command, its reading targets are all tags in a region, e.g., all tags in the 3^{rd} region, so the *Mask* should be set as the region number of the reading targets. For example, in Fig. 8, we set the *Length* $= 4$, *Pointer* $= 0$, and *Mask* to the region number of the 3^{rd} region, i.e., *Mask* $= 0011_2$. In RFace, we let the *security command* and the *collection command* be transmitted alternatively until all tags have been read.

5.1.2 Bit error resistant replay attack detection

During authentication, based on the **efficient reading scheme**, we can get the reading order of tags from the *security command* and collect the backscatter signals of tags from the *collection command*. For detecting replay attacks, we compare the tag reading order specified by security commands with the actual tag reading order. If these two reading orders can not match, we will regard this authentication as a replay attack. As for a $7 * 7$ tag array, we need to transmit nine security commands in total. Therefore, the maximum replay attack success rate can be formulated as $1/49 * 1/40 * 1/31 * 1/22 * 1/13 * 1/10 * 1/7 * 1/4 = 1/4865660800$. It obviously shows that this comparison method performs well in defending against replay attacks.

Although the defense method mentioned above is already feasible in the ideal scene, it is possible to miss reading several tags in reality. In this case, RFace may mistakenly regard a legitimate user as an attacker since the actual tag reading order can not match the preset one perfectly. To avoid such a situation, we allow the tag reading order sequence to have at most 2 bits errors, enabling RFace to be tolerant of bit errors. In this way, for a $7 * 7$ tag array, the maximum replay attack success rate is $1/31 * 1/22 * 1/13 * 1/10 * 1/7 * 1/4 = 1/2482480$. Thus, RFace is still capable to defend against replay attacks with fault tolerance.

5.2 Denoising and Rearrangement

The phase values collected from an RFID reader involve noise due to the involuntary movement of faces and imperfection of hardware. These inevitable defects may make RFace unreliable. We eliminate the noise by unwrapping successive phase values [28]. Moreover, we set a window with a size of five samples and filter out abnormal phase values in the window with the average of other normal values. Fig. 9 shows an instance of noise filtering, where abnormal phase values are filtered out. Additionally, for subsequent facial features extraction, we rearrange both RSS and phase values (which can be represented as $N * 2 * M$,

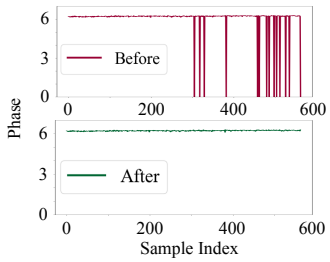


Fig. 9: Performance of denoising method.

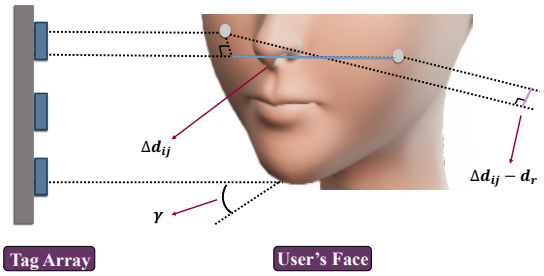


Fig. 10: The impact of deflection.

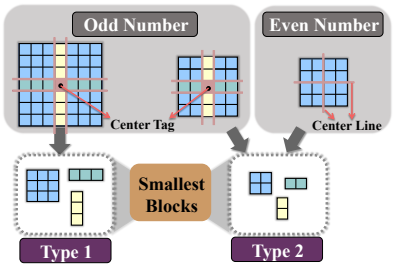
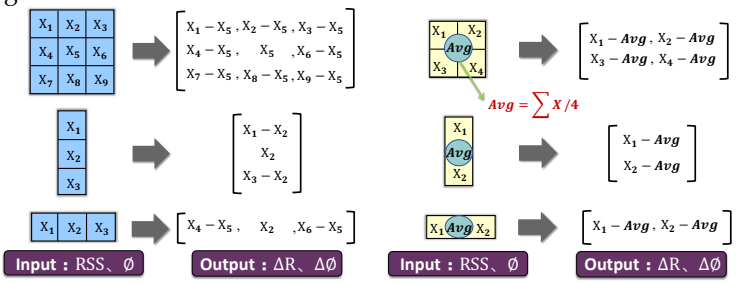


Fig. 11: Division operation in *DR*.



(a) Type1 (b) Type2

Fig. 12: Difference Calculation in *DR*.

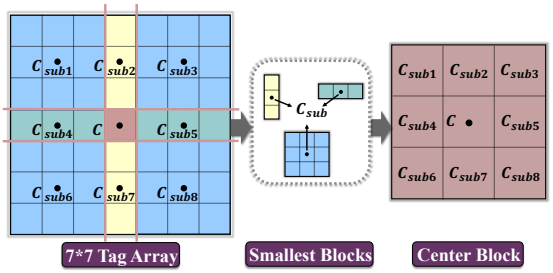


Fig. 13: A 7*7 tag array example represents the processes in *DR*.

N is the number of frames based on slot ALOHA [29], M is the sum of tags) into a new sequence (whose shape is $N * 2 * R * C$, where R is the number of rows and C is the number of columns in tag array) according to the layout of the tag array.

5.3 Disturbance Resistant Feature

Based on the rearranged RSS and phase sequence, we start to extract the 3D geometry and inner biomaterial features (which can be represented as fusion features in the following). As presented in Section 3.1, the RSS difference and phase difference between tags can reveal the fusion features of human faces and we can utilize these two kinds of difference values to extract distance-resistant fusion features. Nevertheless, in practical FA scenarios, it is hard to ensure that the user's face is always in a fixed position and deflection angle relative to the tag array for each authentication attempt. Moreover, due to the sensitivity of RF signals, the tiny variations in distance and angle to the tag array would cause non-trivial errors in the measurements of RSS and phase values [30]. In order to obtain more reliable features, we design a facial feature extraction algorithm resistant to distance and angle variations.

5.3.1 Challenges

Intuitively, we can calculate difference values between any two tags on the tag array to extract desired fusion features. However, we have many practical challenges. Firstly, although arbitrary subtraction can potentially calculate more facial features, it requires about $2 * C_M^2$ times of subtraction for each authentication, which is time-consuming. Additionally, the deflection of the user's face could lead to performance degradation of RFace. As shown in Fig. 10, when a user faces at a certain angle γ, the RSS and phase differences between the leftmost and rightmost tags not only involve the relative distance Δd_{ij}, but also the unexpected distance difference d_r due to facing direction. Δd_{ij} represents the difference of distances between the center points of their corresponding areas to the tag array. This unexpected distance d_r may make RFace mistakenly authenticate a legitimate user as illegitimate. Finally, as shown in Eq. 14, there is an uncertainty term $2k\pi$ which makes the phase difference unstable with the impact of distance variation, thus making RFace unreliable as well.

5.3.2 Disturbance Resistance Algorithm (DR)

In order to tackle the challenges mentioned above, we propose a universal subtraction method that can suppress the disturbance of distance and deflection as well as achieve fast and accurate fusion feature extraction. As for deflection, it is vital to mitigate the unexpected distance difference d_r which may result in authentication deviation. Based on theoretical analysis, we can find that the closer the two tags are selected for subtraction, the smaller the d_r is. Therefore, the key idea of suppressing the impact of face deflection is to narrow the difference calculation range in the tag array. Besides, as for the $2k\pi$ term, the uncertainty of the value k is actually due to the variation of distance. Here, we can eliminate this uncertainty by calculating the cosine of the phase difference to obtain stable distance-resistant features.

According to the above analysis, we propose the disturbance resistance (DR) algorithm (as shown in Algorithm 1) to extract reliable fusion features. It is worth noting that the DR algorithm can be flexibly applied to square tag arrays of any size. The purpose of DR is to decrease the time for difference calculation and avoid the RSS subtraction and phase subtraction for a pair of tags that are far apart on the tag array (*e.g.*, the leftmost and rightmost, top and bottom). The DR algorithm can be divided into three steps: division, difference calculation, and reorganization.

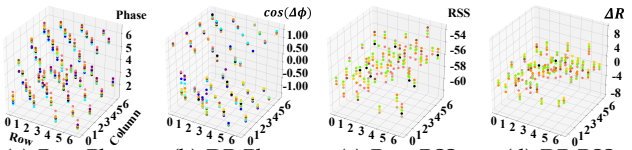
Division: We first apply the division operation (Division function in Algorithm 1) to find the center of the tag array and divide it into smaller blocks. For the square tag arrays, their layouts can be divided into two categories: the number of rows is odd or even. Different methods for the division

Algorithm 1: Disturbance Resistance

Input: 3-dimension matrix M with RSS and θ of each tag;
Output: 2-dimension fusion feature array:
$\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\}, r \in [1, M.row], c \in [1, M.column]$;

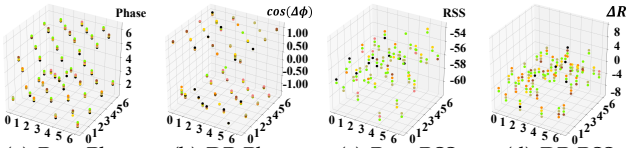
DR(M);
Function `DR(`M`)`**:**
  if M *is smallest block* **then**
    $\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\} \leftarrow$ **Calculation(M);**
  else
    $\{B, C_{set}\} \leftarrow$ **Division(M) ;**
    for *each smallest block b in B* **do**
      $\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\} \leftarrow$ **Calculation(b);**
    end
    if C_{set} *is not empty* **then**
      $C_{block} \leftarrow$ **Reorganization(C_{set});**
      $\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\} \leftarrow$ **DR(C_{block});**
    end
  end
  return $\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\}$;
Function `Calculation(`b`)`**:**
  if $b \in Type1$ **then**
    $RSS_{Center} \leftarrow RSS_{\frac{b.row+1}{2}, \frac{b.column+1}{2}}$;
    $\theta_{Center} \leftarrow \theta_{\frac{b.row+1}{2}, \frac{b.column+1}{2}}$;
  else if $b \in Type2$ **then**
    $RSS_{Center} \leftarrow$ **Average(b, RSS)**;
    $\theta_{Center} \leftarrow$ **Average(b,** θ**)**;
  end
  for $r \leftarrow [1 : b.row]$ **do**
    for $c \leftarrow [1 : b.column]$ **do**
      if $r == \frac{b.row+1}{2}$ *and* $c == \frac{b.column+1}{2}$ **then**
        Continue;
      else
        $\Delta R_{r,c} \leftarrow RSS_{r,c} - RSS_{Center}$;
        $cos(\Delta\phi_{r,c}) \leftarrow cos(\theta_{r,c} - \theta_{Center})$;
      end
    end
  end
  return $\{\{\Delta R_{r,c}\}, \{cos(\Delta\phi_{r,c})\}\}$;
Function `Division(`M`)`**:**
  if $M.row$ *or* $M.column$ *is odd number* **then**
    Add the center point of M **into** C_{set};
    Divide M **along center tag into smaller blocks** B_1;
  end
  else
    Divide M **along center lines into smaller blocks** B_1;
  end
  for *each block b in* B_1 **do**
    if *b is smallest block* **then**
      Add b **into** B;
    else
      $\{B', C'\} \leftarrow$ **Division(b);**
      Add B' **into B, Add** C' **into** C_{set};
    end
  end
  return $\{B, C_{set}\}$;

of these two kinds of tag arrays are shown in Fig. 11. As for the tag array with odd rows, we regard the center tag of the tag array as the center point and add it into the center point set (C_{set} in Algorithm 1). Then, we divide the tag array into smaller blocks (B_1 in Algorithm 1) based on the row and column where the center tag locates. As for the tag array with even rows, there is no center tag. Thus, we divide the tag array into smaller blocks (B_1) along the center lines of the tag array. The center lines represent the lines connecting the midpoints of opposite sides of the square tag array. Then, we repeat this division operation for these smaller blocks (B_1) until we obtain the smallest blocks (b in Algorithm 1) whose rows and columns are both no more than 3. As shown in Fig. 11, the smallest blocks (b) can be



(a) Raw Phase (b) DR Phase (c) Raw RSS (d) DR RSS

Fig. 14: Distributions of raw RSS and phase and features extracted by DR in different distance conditions.



(a) Raw Phase (b) DR Phase (c) Raw RSS (d) DR RSS

Fig. 15: Distributions of raw RSS and phase and features extracted by DR in different deflection conditions.

divided into two types. Type 1 denotes the tag blocks whose number of rows or columns is three, i.e., the shape of the block is $1*3$, $3*1$, or $3*3$. Type 2 represents the tag blocks whose number of rows or columns is two, i.e., the shape of the block is $1*2$, $2*1$, or $2*2$.

Difference Calculation: After the division operation, for each obtained smallest block, we start to calculate the RSS difference and the phase difference (Calculation function in Algorithm 1) to extract the stable fusion features. As shown in Fig. 12(a), If the smallest block belongs to Type 1, we will first find the center tag in the block. Then, we calculate the RSS difference (ΔR) and the phase difference ($cos(\Delta\phi)$) between each remaining tag and the center tag. If the smallest block belongs to Type 2, we will first calculate the average values of the RSS and phase for this smallest block. Then, we calculate the RSS difference (ΔR) and the phase difference ($cos(\Delta\phi)$) between each tag on the smallest block and the average values (as shown in Fig. 12(b)).

Reorganization: Finally, for the unassigned center point set (C_{set}), we reorganize it into a center block (C_{block} in Algorithm 1) according to the original position of center tags (Reorganization function in Algorithm 1). Taking a $7*7$ tag array as an example, Fig. 13 shows the reorganization process, i.e., from $7*7$ tag array to the center block. Then, we apply the DR algorithm for the center block (C_{block}) to get a complete fusion feature array.

5.3.3 Feasibility of DR

In order to show the feasibility of DR, we use a 7*7 tag array as an example and show the effect of our DR algorithm.

As shown in Fig. 13, we first find the center point (C) of the tag array and divide the tag array into smaller blocks based on the row and column where C locates. Then, we repeat this division operation until we obtain the smallest blocks (belong to Type 1). Next, for each smallest block, we first find the center of it (C_{sub}), and then for each remaining tag (represented by $T_{r,c}$, with r and c as the row and column it locates) in this block, we get:

$$\Delta R_{r,c} = RSS_{r,c} - RSS_{C\_sub}, \tag{15}$$

$$\Delta\phi_{r,c} = \theta_{r,c} - \theta_{C\_sub}, \tag{16}$$

with $\Delta R_{r,c}$ and $\Delta\phi_{r,c}$ denoting the RSS difference and phase difference between $T_{r,c}$ and the corresponding center point C_{sub} of the block. For the unassigned center points (C_{sub}) of these smallest blocks, we reorganize them into a

center block. Then we apply the DR algorithm to the center block to calculate the RSS difference and phase difference between the C_{sub} and the center point of the whole tag array (C). So far, we have obtained ΔR and $\Delta\phi$ of all tags. As analyzed at the beginning, we have to substitute $\Delta\phi$ by $cos(\phi)$ in the phase part of our final fusion features. Finally, for one frame of RFID signals, combining the ΔR and $cos(\Delta\phi)$ of all tags in the tag array, we can form a 3-dimension fusion feature array with a shape of $(2 * R * C)$ with only $(2 * R * C)$ times of subtraction operations.

Fig. 14 and Fig. 15 show the effect of our DR algorithm with the 7*7 tag array, where the X-axis and Y-axis represent the row and column of the tag array respectively. Compared with the distribution of raw RSS and phase of each tag before (Fig. 14(a), Fig. 14(c) and Fig. 15(a), Fig. 15(c)), the values of ΔR and $cos(\Delta\phi)$ after applying the DR algorithm (Fig. 14(b), Fig. 14(d) and Fig. 15(b), Fig. 15(d)) become more stable under varying distance and deflection conditions. Hence, we can obtain reliable fusion features based on the DR algorithm.

5.4 Anti-Spoofing Authentication

Based on the extracted facial fusion features, we next demonstrate how to perform face authentication and defend against spoofing attacks. In addition to the aforementioned RSS difference and phase difference, we also include the time series as a new dimension in the feature extraction. We take A ($A = 5$ in implementation) consecutive fusion feature arrays $(2 * R * C)$ in time series as a fusion feature block $(A * 2 * R * C)$, and reshape the fusion features which we obtain from DR algorithm into a five-dimensional array $(N/A * A * 2 * R * C)$. Then, the following anti-spoofing authentication and feature granularity exploration are all based on this five-dimensional fusion feature array.

In RFace, we utilize an SVM as the classifier to implement face authentication and spoofing attack detection. Specifically, each user needs to provide a batch of signal samples to train the SVM for RFace registration. All feature blocks are reshaped as one-dimensional feature vectors to train the SVM. In the authentication stage, once receiving a login request, RFace feeds extracted features from the received signal samples into the pre-trained SVM model. Then, it outputs a series of confidence coefficients representing the similarity between the login user and the registered users in the database. RFace finds the largest one among these confidence coefficients and compares it with a predetermined threshold. If it is larger than the threshold, the user will be accepted as a legitimate user. Otherwise, the user will be denied. Since the fusion features of each person are distinct, unregistered users will fail the authentication owing to low similarity. Furthermore, as spoofing attackers cannot produce inner biomaterial features, attackers will be rejected by the threshold-confidence comparison mechanism.

6 Evaluation

We evaluate the performance of RFace in real environment on both authentication and anti-spoofing.

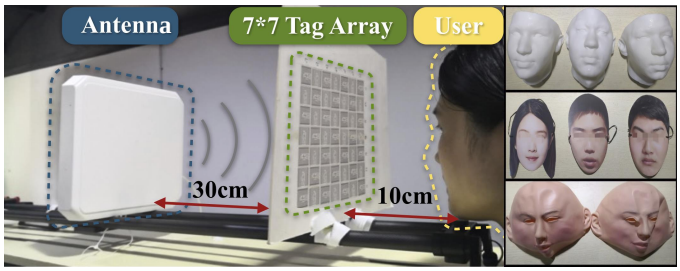


Fig. 16: Experimental setup.

6.1 Implementation

Hardware: RFace is implemented with an *ImpinJ R420* reader connected to a directional antenna *Larid A9028*. Besides, we utilize 49 *Alien-9629* RFID tags to form a $7 * 7$ tag array in a perpendicular orientation deployment as shown in Section 2. The working frequency of RFace is $922.38MHz$.

Software: The collection of RF signals adopts EPC Gen2 standard to avoid collision and control the communication respectively. The client communicates with reader based on the ImpinJ LLRP Tool Kit (LTK). It delivers Gen2 parameters (e.g., *Pointer*, *Mask*, *Length*) to specify the operation of reader, i.e., selectively reading tags. The processing algorithms of RFace are implemented by *Python* which runs on a personal computer (PC) with Intel(R) Core(TM) *i5-8250U* 1.6 GHz CPU and 8 GB RAM.

Experimental Setup: We conduct our experiments in three real environments: seminar room, lab, and office. Fig. 16 shows the default setup of RFace. The directional antenna is placed $30cm$ behind the tag array to transmit interrogative RF signals, and users place their faces around $10cm$ in front of the tag array for authentication. Then RF signals reflected from human faces are collected and further sent to the PC via Ethernet for processing.

Data Collection: We conduct the experiments by adhering to the approval of our university's Institutional Review Board (IRB). We invite 30 volunteers (10 females and 20 males) aged from 18 to 30 in our experiments. Among the 30 volunteers, we randomly choose 5 volunteers as spoofers and the rest 25 volunteers register as legitimate users. In the registration phase, we collect three groups of RF signals (each group contains 60 fusion feature blocks) for each legitimate user, which takes about 225 seconds. Then in the authentication phase, each legitimate volunteer performs authentication for 120 times and each illegitimate volunteer performs 360 times authentication to test the performance of RFace. Each authentication attempt takes about 1.25 seconds.

Metrics: We define six metrics to evaluate RFace: Authentication Success Rate (ASR), False Accept Rate (FAR), False Reject Rate (FRR), Receiver Operating Characteristic (ROC), Equal Error Rate (EER) and Defense Success Rate (DSR). ASR is the probability that the system authenticates a legitimate user correctly and can be represented as: $ASR = \frac{N_{acc}}{N_all}$, where N_{acc} is the number of correct authentication times for legitimate users and N_{all} is the number of all authentication times for legitimate users. FAR is the probability that the system mistakenly authenticates an illegitimate user as a legitimate one and can be represented as: $FAR = \frac{N_{wr}}{N_{il}}$, where N_{wr} is the number of wrong

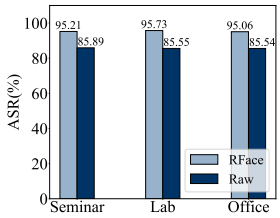


Fig. 17: ASR under different environments.

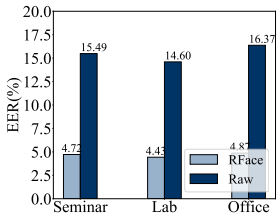


Fig. 18: EER under different environments.

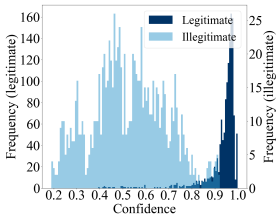


Fig. 19: Confidence distributions for users.

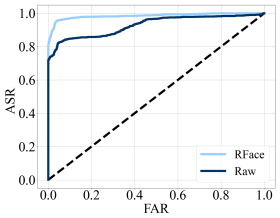


Fig. 20: ROC curves for RFace and raw data.

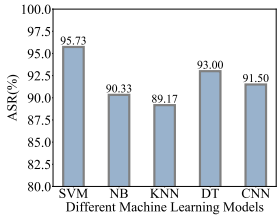


Fig. 21: ASR for different machine learning models.

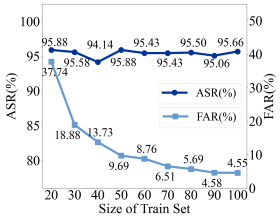


Fig. 22: The impact of the training set size.

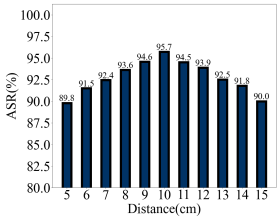


Fig. 23: The impact of distance between user and tag array on ASR.

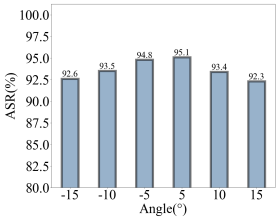


Fig. 24: The impact of deflection on ASR.

accepted times for illegitimate users and N_{il} is the number of all authentication times for illegitimate users. FRR is the probability that the system mistakenly authenticates a legitimate user as an illegitimate one. ROC curve indicates the relationship between the ASR and FAR under various threshold. Additionally EER describes the rate where FAR equals FRR. Finally, we define the DSR to measure the performance of RFace on anti-spoofing. DSR is the probability that a spoofing attack is successfully detected, and the higher DSR indicates that RFace is more secure. DSR can be formulated by: $DSR = \frac{N_{def}}{N_{att}}$, where N_{def} is the number of successfully detected spoofing attacks and N_{att} is the number of all spoofing attacks.

6.2 Overall Performance

We evaluate the overall performance of RFace in three different environments. We first compare the overall performance of RFace with/without our DR algorithm. As shown in Fig. 17, the average ASRs for RFace and raw data are 95.3% and 85.6%, respectively. Meanwhile, with the threshold of 0.8, as shown in Fig. 18, the average EERs of RFace and raw data are 4.67% and 15.49%, respectively. These comparison results demonstrate that our denoising and feature extraction method can effectively remove noise and improve feature quality. To show if RFace is able to reject illegitimate users, we plot the confidence distributions of legitimate users and illegitimate users in Fig. 19. This result indicates that RFace can effectively reject illegitimate users with a high probability. Therefore, as shown in Fig. 20, RFace can achieve a low FAR of 4.48% while the FAR of raw data is as high as 16.52%. These results demonstrate that RFace can authenticate users accurately and securely. In addition, our experiments show that every authentication only takes 1.26 seconds on average, indicating that RFace performs well in real-time authentication.

6.3 Comparison with Mainstream Classifiers

To show the superiority of our selection of classifier (i.e., SVM), we compare the performance of SVM with different machine learning models. First, we compare SVM with other three mainstream and traditional classifiers: Naive Bayes (NB), K-Nearest Neighbors (KNN), and Decision Tree (DT). As shown in Fig. 21, the ASRs for SVM, NB, KNN, and DT are 95.7%, 90.3%, 89.2%, and 93.0%, respectively. These results indicate that SVM outperforms these classifiers. Meanwhile, the ASRs of the other three classifiers are also high, demonstrating that the feature extraction method in RFace is significantly effective. Additionally, we also compare the performance of RFace with a typical Convolutional Neural Network (CNN) [31]. With the same data set, the ASR and FAR of CNN are 91.5% and 15%, respectively. Such results indicate that SVM performs better than CNN when the data set is small or disturbed by environmental noise. Besides, training a CNN model requires more computation than SVM, e.g., a complete CNN training process takes 80 seconds while the SVM only takes 3 seconds. Therefore, considering the accuracy and costs, we finally choose SVM as our classifier.

6.4 Impact of Training Set Size

To explore the impact of training set size, *i.e.*, the number of feature blocks of each user in the training set, we vary the size from 20 to 90. The experiment results shown in Fig. 22 indicate that with 90 feature blocks for each user, RFace can achieve a FAR that is lower than 5.0% while retaining a high ASR of about 95.5%. This result also shows that RFace is user-friendly in user registration because collecting 90 feature blocks only takes 112.5 seconds approximately.

6.5 Performance Towards Distance and Deflection Variations

In the distance experiment, we use the feature blocks collected when the user's face is placed in front of the tag array 10 centimeters as the training set. Then we vary the distance between the tag array and the user from 5 centimeters to 15 centimeters to evaluate the impact of distance. The authentication results are shown in Fig. 23, in which we find that large distance variation (*i.e.*, far from 10 centimeters) would cause ASR reduction. However, even if the difference is 5 centimeters, the reduction scale is still acceptable (less

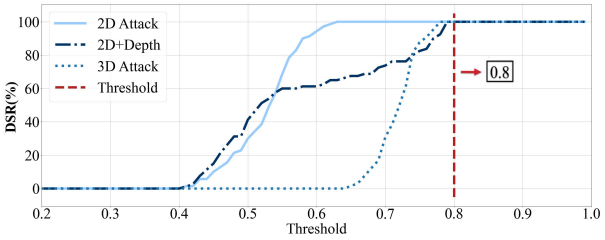
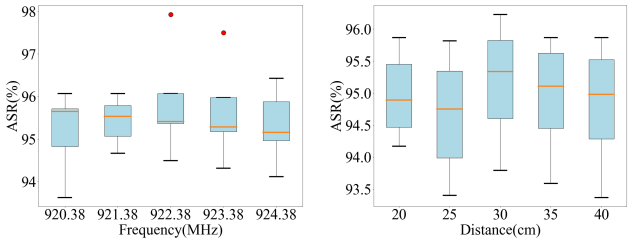


Fig. 25: The DSR of three attacks under various thresholds.



(a) Frequency (b) Distance

Fig. 26: The impact of working frequency and distance between antenna and tag array on ASR.

than 6%), which proves that the DR algorithm is effective at distance disturbance suppression and RFace is robust to distance variation.

In the deflection experiment, we use the normal feature blocks without deflection as the training data. Then we collected testing data of two orientations (*i.e.*, leaning cheeks to left or right) of deflections. Specifically, we vary the deflection angles from 5 degrees to 15 degrees. As in Fig. 24, it can be observed that even if the left or right deflection is 15 degrees, RFace can still achieve a high ASR. Therefore, DR is also effective in deflection impact suppression.

6.6 Attack and Defense

Attack Realizations: We evaluate our system against the replay attack and three types of spoofing attacks: 2D spoofing attack, 2D+ spoofing attack and 3D spoofing attack. For the replay attack, we randomly select one RF signal from the previous successful authentication data to deceive RFace. We repeat the replay attack 100 times to evaluate the security of RFace. For the 2D spoofing attack, there are two attack methods including static and dynamic attacks, which are respectively realized with photos and videos of three victims (each authenticates 60 times) in front of the tag array. For the 2D+ spoofing attack, we recruited 15 unregistered volunteers as attackers, each wearing a soft PVC mask printed with the victim's photo respectively. We make three soft PVC masks with three victims' photos and each attacker wears each mask to attack RFace 60 times. For the most challenging 3D spoofing attack, we launch attacks by utilizing photosensitive resin to build three 1 : 1 scale 3D masks of victims, which simulate both 2D features and precise depth features of the victim. Specifically, we also pose each 3D printed mask in front of the tag array to perform 3D spoofing attack 60 times. The setup of these three types of attack is presented in Fig. 16.

TABLE 1: Comparison with previous work

System	2D Attack Resistance	2D+ Attack Resistance	3D Attack Resistance	Privacy Preserved
Samsung FR [32]	×	×	×	×
EchoFace [33]	✓	✓	×	×
FaceHeart [34]	✓	✓	×	×
Face Flashing[3]	✓	✓	×	×
RFace	✓	✓	✓	✓

Defense Performance: We calculate the DSRs of the above attacks. The DSR of the replay attack is 100% which indicates that no replay attack succeeds in deceiving RFace in the experiments. The high DSR demonstrates that RFace can effectively defend against the replay attack. The spoofing attack results are shown in Fig. 25, which indicates that the DSRs for all three types of attacks are 100%. This is because the confidence coefficients of the features provided by attackers cannot reach our empirical threshold of 0.8. The results indicate that the extracted inner biomaterial feature is effective in defending against spoofing attacks. Additionally, in order to further verify RFace's ability to defend against spoofing attacks, we conduct the poisoning attack experiment. Specifically, in the poisoning attack, a malicious user wears a soft PVC mask for registration, and then a group of attackers wear the same soft PVC mask for authentication. In this way, attackers can roughly simulate both 2D structure features and biomaterial features of the maliciously registered user. To evaluate RFace's defense ability against this attack, we add the RF signals of a user wearing a soft PVC mask to the training data. Then, we ask ten users wearing the same soft PVC mask to attack RFace. It is worth noting that in the poisoning experiment, no attacker succeeds in deceiving RFace. The reason is likely to be the following two aspects. First, the 3D facial structures of different users are various, so even if they wear the same soft PVC mask, the 3D geometry features extracted by RFace are inconsistent. Second, RF signals have excellent penetrability, so the biomaterial features captured by RFace should include both the soft PVC mask and the attacker's own facial biomaterial features. Therefore, even wearing the same soft PVC mask, RFace can still distinguish the attacker. Table 1 compares RFace with four advanced FA systems in terms of spoofing attack defense ability. The results show that, compared with other FA systems, RFace can protect users' privacy and defend against various spoofing attacks.

6.7 Impact of Experimental Setup

To explore the impacts of different experiment setups, we vary the working frequency and the distance between the antenna and the tag array respectively, and observe the variation of authentication performance.

Firstly, we vary the working frequency of RFace from 920.38MHz to 924.38MHz with the step of 1Mz. The experiment results are shown in Fig. 26(a). As we can observe, the ASRs of different frequencies are very similar and they are all higher than 95%. These results indicate that the variation of working frequency has negligible impacts on RFace's authentication performance. By setting different working

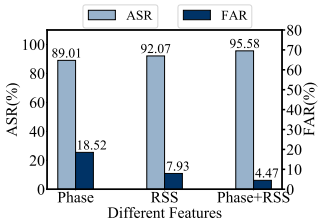


Fig. 27: The impact of different features on FA performance.

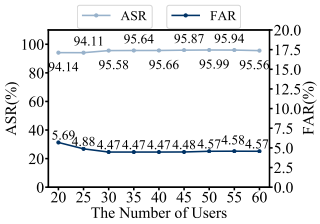


Fig. 28: The impact of user number on FA performance.

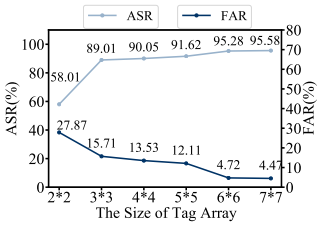


Fig. 29: The impact of tag array size on FA performance.

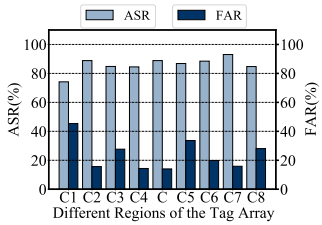


Fig. 30: The impact of different regions of the tag array on FA performance.

frequencies, multiple RFace systems can be deployed in the same location without interference. Then, we vary the distance between the antenna and the tag array from 20cm to 40cm with a step of 5cm. The experiment results are shown in Fig. 26(b), from which we can find the mean ASRs are all around 95%. So, we can conclude that the distance variation (within $[20, 40]cm$) has negligible impacts on authentication performance. Thus, users can flexibly deploy RFace according to their demands.

6.8 Impact of Different Features

RFace extracts fusion facial features from both phase and RSS data for facial authentication. To show the FA performance when using different features, we utilize phase difference, RSS difference and the combination of them to conduct facial authentication, respectively. The experiment results are shown in Fig. 27. We can find that when only phase difference or RSS difference is used, the ASR is 89% and 92%, respectively. The ASRs in these two cases are lower than the ASR (95.6%) when using the combination of phase difference and RSS difference. Moreover, the FAR of using the combination of phase difference and RSS difference for FA is 4.47%, which is lower than using only RSS or phase for FA. It indicates that utilizing the combination of RSS and phase can provide more facial characteristics and achieve more reliable FA.

6.9 Impact of User Number

To investigate whether RFace can support a larger user set, similar to the experimental methodology in [35], we explore how the performance of RFace changes as the number of users increases. To this end, we increase the number of users from 20 to 60 and calculate the corresponding FA accuracy. As shown in Fig. 28, when the number of users reaches 60, ASR is still higher than 95.5% and FAR is lower than 4.6%. It indicates that the performance of RFace does not change significantly as the user number increases. It also provides encouraging signs that RFace has good scalability and can be well applied in the access control of group clients.

6.10 Impact of Tag Array Size

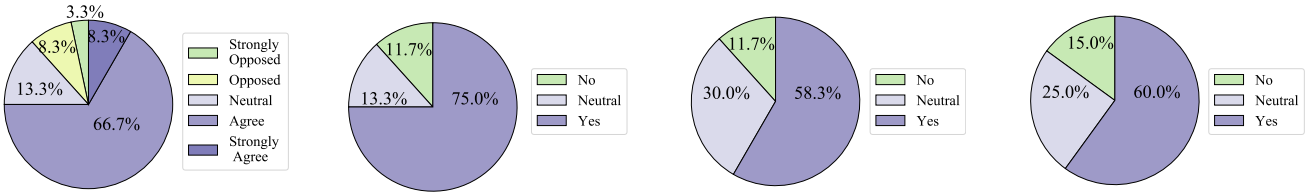
To evaluate the impact of tag array size on FA performance, we increase the number of tags from $2 * 2$ to $7 * 7$ and analyze the changing trend of FA accuracy. As shown in Fig. 29, we can find that the larger the tag array, the higher the FA accuracy. When the size of the tag array is $7 * 7$, RFace achieves the best performance with a high ASR of 95.6% and a low FAR of 4.47%. Moreover, we can observe that when the tag array size is larger than $6 * 6$, the ASR is already higher than 95% and the FAR is lower than 5%. This is because the $6 * 6$ tag array is capable to cover most of the human face, so it can capture enough facial features to support FA. Therefore, when choosing the tag array for FA, we recommend selecting the one whose size is similar to the size of the human face. The reason is that when the tag array is too big, it will not only capture the signals reflected by the human face but also the signals reflected by other irrelevant objects, thus resulting in a degradation in authentication accuracy.

6.11 Impact of Different Regions of the Tag Array

In order to explore the impact of different regions of the tag array on feature extraction and further understand the signal propagation theoretical model (Section 3), we conduct experiments by utilizing the features captured by different regions on the tag array for FA and calculate corresponding accuracy. Specifically, taking the default $7 * 7$ tag array as an example, we divide the tag array into nine regions, and each region is a $3 * 3$ small tag array centered on the center tag $c_{sub1}...c...c_{sub8}$ respectively, as shown in Fig. 13. We denote the $3 * 3$ small tag arrays centered on $c_{sub1}...c...c_{sub8}$ as $C1...C...C8$, respectively. Then, we utilize the features captured by the regions $C1...C...C8$ for FA. The results are shown in Fig. 30. We can observe that the regions $C2$, C, and $C7$ have higher ASR and lower FAR than other regions. This is most likely because these regions capture the signals mainly reflected by the middle part of the human face, which contains more face-related features. As illustrated in Section 3, different tags can be mapped to different areas on the human face, so the contribution of features captured by different tags to FA depends on the discriminability of their corresponding face area.

6.12 User Study

We conduct a series of questionnaire surveys of 60 users to evaluate the usability of RFace. Fig. 31 shows the specific questions in the questionnaire and the corresponding results. Fig. 31(a) indicates that 88.4% of users think that RFace offers additional advantages over vision-based FA methods, e.g., privacy-preserving, anti-spoofing, and robustness to light conditions. Additionally, 75% of users think RFace is convenient and only 11.7% of users feel uncomfortable when using RFace. As for the applicability,



(a) The percentage of users who believe that RFace offers additional advantages over vision-based FA methods

(b) The percentage of users who think RFace is convenient

(c) The percentage of users who feel comfortable using RFace

(d) The percentage of users who think it is acceptable to use RFace for FA daily in the workplace

Fig. 31: Usability study for RFace.

only 15% of users are unwilling to use RFace for FA daily in the workplace. From these usability studies, we can find that RFace is user-friendly and has the potential to complement and enhance existing FA systems in some challenging scenarios.

7 RELATED WORK

7.1 Facial Authentication

FA has been extensively studied in both literature and industrial communities. Existing FA systems can be divided into vision-based ones [23, 36] and wireless signal-based ones [37, 38, 39]. Vision-based approaches use RGB cameras to collect facial information. They can achieve significantly high accuracy and decent robustness. However, they also have two major limitations, namely the risk of privacy leakage [7, 8] and vulnerability to spoofing attacks [23]. First, in a camera-based FA service, the untrusted third-party server may abuse users' data, e.g., maliciously link users' face images to other sensitive data like health records, raising privacy leakage concerns [8, 12]. Although researchers have proposed some privacy-preserving solutions [9, 10, 11, 12], e.g., using cryptographic protocol [9, 10, 11] and adding differential privacy-based perturbation to original face images [12]. However, such privacy issues are not solved fundamentally. Any untrusted third parties (e.g., the camera manager) who own the original images still could leak users' privacy [12, 40]. On the other hand, traditional camera-based FA systems are vulnerable to spoofing attacks. They could be tricked by photos or videos [1, 3, 23, 38]. To address this problem, some advanced FA systems (e.g., Apple FaceID [41]) probe the 3D facial information of the user before making authentication decisions. Nevertheless, these systems still could be deceived by manipulated 3D inputs (e.g., 3D printed face masks [42]).

To tackle these limitations, wireless-based FA systems are designed to utilize the human-eye imperceptible wireless signal, e.g., acoustic [38], millimeter wave (mmWave) [39] and RFID [37], for facial authentication. For instance, EchoPrint [38] leverages acoustic signals to capture the facial features and determine whether the object in front of the sensor is a 3D one. Such an acoustic-based FA system can protect users' visual privacy as well as resist 2D spoofing attacks. However, it can not defend against the more advanced 3D spoofing attacks. mmFace [39] utilizes the mmWave signals to capture both the facial biometric features and structure features, making it resilient to 2D and 3D spoofing attacks. Nevertheless, its implementation requires costly mmWave radar. Compared with mmWave and acoustic signals, RFID bears many advantages in implementing FA. As another work using RFID for FA, RFaceID [37] also employs passive RFID arrays to capture facial features. Different from our work, RFaceID directly constructs a deep neural network to extract facial features instead of building a theoretical model like RFace to probe the correlation between facial features and RF signals. Besides, RFaceID requires the users to shake their faces during FA and there is no strategy designed to defend against signal replay attacks. RFaceID also did not measure its resistance to 2D, 2D+ and 3D spoofing attacks. Comparatively, RFace does not suffer from these flaws. It builds a privacy-preserving FA system with a low-cost RFID tag array. RFace does not require user interaction and meanwhile defends against both spoofing attacks (2D, 2D+ and 3D) and signal replay attacks.

7.2 RFID-based User Authentication

RFID has been developed to authenticate users' identities in recent years [43, 44]. Current RFID-based user authentication approaches can be categorized into two groups according to whether the user needs to touch the tags or not: contact-based [28, 45, 46] and contact-free ones [47, 48]. In contact-based authentication systems, users are required to touch the tags in a tag array. In this process, users' biomaterial features (e.g., impedance [28], geometry [45], and composition [45]), and behavioral features like touching patterns and drawing habits [45] are embedded in the signals and backscattered to the reader. As the user needs to either carry a tag array or cooperatively interact with the system (e.g., tapping with a specific rhythm [49]), these approaches would bring inconvenience to users. By contrast, contact-free methods usually require fewer interactions between the users and the systems. To collect biomaterial features, users only need to exhibit their body parts (e.g., hand [47]) to the tag array. In this paper, we also seek for minimizing the requirements for the users and propose RFace. In addition to the user-friendly authentication manner, the security of RFace is also greatly enhanced by the proposed replay defense method (section 5.1). Compared with existing replay defense approaches [20, 47], our approach bears two major advantages: (1) Our approach does not modify the original signals, while that in [20] adds elaborate noise in emitted signals, which would destroy the biometrics contained in the signals. (2) Our approach does not introduce extra storage overhead, but that in [47] must store pre-used EPC sequences in the database.

8 Discussion

Practicality: (1). Accuracy: The experiment results show that RFace achieves 95.7% authentication success rates and 4% equal error rates. Although the FA accuracy of RFace is better than another state-of-the-art RFID-based FA system [37], we admit that the accuracy of RFace is not comparable to current advanced camera-based FA systems. However, we want to clarify that RFace is not designed to substitute the camera-based methods, but instead shows the potential to complement and enhance existing FA systems in some challenging scenarios, such as suffering from spoofing attacks or privacy leakage. In practice, to increase the authentication accuracy, if a user fails the authentication, RFace can conduct a second-time authentication for the user. Increasing the number of authentication attempts can effectively reduce the false reject rate and the time required for each authentication is really short. (2). The number of users: The current version of RFace is designed for the access control of group clients in public institutions (e.g., communities, small- or moderate-scale companies) or individual identification of smart home devices. In these scenarios, the number of users would not be as large as that of mobile applications, which may be in a scale of millions. Moreover, according to the evaluation results, RFace performs well when the amount of users is on a moderate-scale.

Robustness: (1). Authentication distance: To ensure a high authentication accuracy, we recommend that users pose their faces close to the tag array. We believe such a requirement on sensing distance can be satisfied in practical authentication applications, e.g., access control at entrances and individual identification of the smart home device. (2). Obstacle: Like existing FA systems, the performance of RFace would degrade if the line-of-sight (LOS) path between the tag array and face is obstructed. However, it is very rare that there are obstacles blocking the LOS path in reality. Besides, obstacles can be easily detected and users can remove them before authentication. (3). Interference: Due to the openness of the physical world, the RF signals emitted by the reader may be interfered by other RF signal sources. However, the interference only functions when these two kinds of signals are at the same frequency band, while the devices operating under the permitted RFID frequency range are not common in our daily life. Moreover, thanks to this property, users can deploy multiple RFace systems working at different frequency bands in the same place to improve the authentication efficiency.

9 Conclusions

In this paper, we build a novel facial authentication system named RFace with COTS RFID devices. RFace ensures privacy-preserving and spoofing-resistant simultaneously. We build a rigorous theoretical model to prove the feasibility of extracting both 3D geometry and biomaterial features from backscatter RFID signals. To enhance the security of our system, we design a replay attack defense method by specifying the tag reading order. In order to alleviate the impact caused by position differences in real scenarios, we design a novel disturbance resistance algorithm. We conduct comprehensive evaluations with 30 participants in various experiment settings. The results show that RFace achieves an ASR of 95.7% and an EER of 4.4%. More importantly, RFace can effectively defeat spoofing attacks by jointly considering 3D geometry and biomaterial features.

Acknowledgments

This paper is partially supported by the National Key R&D Program of China (2021QY0703), National Natural Science Foundation of China under grant U21A20462, 61972348, Research Institute of Cyberspace Governance in Zhejiang University, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), Basic Research Project of Shenzhen Science and Technology Innovation Commission (Project Number: JCYJ20190812155213250), Shenzhen Longhua District Science and Technology innovation special fund project (Project Number: JCYJ201903), the National Science Fund for Distinguished Young Scholars of China under grant No.62125203, Hong Kong General Research Fund (GRF) under grant PolyU 152165/19E.

References

[1] H. Chen, W. Wang, J. Zhang, and Q. Zhang, "Echoface: Acoustic sensor-based media attack detection for face authentication," *IEEE Internet of Things Journal (IoTJ)*, vol. 7, no. 3, pp. 2152–2159, 2020.

[2] B. Zhou, J. Lohokare, R. Gao, and F. Ye, "Echoprint: Two-factor authentication using acoustics and vision on smartphones," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[3] D. Tang, Z. Zhou, Y. Zhang, and K. Zhang, "Face flashing: a secure liveness detection protocol based on light reflections," in *Proceedings of the Annual Network and Distributed System Security Symposium (NDSS)*, 2018.

[4] S. Rajarajan and P. Priyadarsini, "UTP: a novel PIN number based user authentication scheme," *International Arab Journal of Information Technology*, vol. 16, no. 5, pp. 904–913, 2019.

[5] A. S. Rathore, W. Zhu, A. Daiyan, C. Xu, K. Wang, F. Lin, K. Ren, and W. Xu, "Sonicprint: a generally adoptable and secure fingerprint biometrics in smart devices," in *Proceedings of the ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2020.

[6] S. Mare, M. Baker, and J. Gummeson, "A study of authentication in daily life," in *Proceedings of the USENIX Symposium on Usable Privacy and Security (SOUPS)*, 2016.

[7] A. Bhargav-Spantzel, A. Squicciarini, and E. Bertino, "Privacy preserving multi-factor authentication with biometrics," in *Proceedings of the second ACM workshop on Digital identity management*, 2006, pp. 63–72.

[8] J. Bringer, H. Chabanne, and A. Patey, "Privacy-preserving biometric identification using secure multiparty computation: An overview and recent trends," *IEEE Signal Processing Magazine*, vol. 30, no. 2, pp. 42–52, 2013.

[9] A.-R. Sadeghi, T. Schneider, and I. Wehrenberg, "Efficient privacy-preserving face recognition," in *International Conference on Information Security and Cryptology*. Springer, 2009, pp. 229–244.

[10] P. Terhörst, K. Riehl, N. Damer, P. Rot, B. Bortolato, F. Kirchbuchner, V. Struc, and A. Kuijper, "Pe-miu: A training-free

privacy-enhancing face recognition approach based on minimum information units," *IEEE Access*, vol. 8, pp. 93 635–93 647, 2020.

[11] Z. Ma, Y. Liu, X. Liu, J. Ma, and K. Ren, "Lightweight privacy-preserving ensemble classification for face recognition," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5778–5790, 2019.

[12] M. A. P. Chamikara, P. Bertók, I. Khalil, D. Liu, and S. Camtepe, "Privacy preserving face recognition utilizing differential privacy," *Computer & Security*, vol. 97, p. 101951, 2020.

[13] R. AMADEO, "Galaxy s8 face recognition already defeated with a simple picture," https://arstechnica.com/gadgets/2017/03/video-shows-galaxy-s8-face-recognition-can-be-defeated-with-a-picture/, 2017.

[14] K. Kollreider, H. Fronthaler, and J. Bigun, "Non-intrusive liveness detection by face images," *Journal of Image and Vision Computing*, vol. 27, no. 3, pp. 233–244, 2009.

[15] A. Wyatt, "What exactly is the dot projector? why it is used in iphone x?" https://www.thebestintech.com/what-is-dot-projector/.

[16] N. Erdogmus and S. Marcel, "Spoofing face recognition with 3d masks," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 9, no. 7, pp. 1084–1097, 2014.

[17] D. Vasisht, G. Zhang, O. Abari, H. Lu, J. Flanz, and D. Katabi, "In-body backscatter communication and localization," in *Proceedings of the International Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, 2018.

[18] L. Yang, Q. Lin, X. Li, T. Liu, and Y. Liu, "See through walls with cots rfid system!" in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2015.

[19] C. Wang, J. Liu, Y. Chen, H. Liu, L. Xie, W. Wang, B. He, and S. Lu, "Multi-touch in the air: Device-free finger tracking and gesture recognition via COTS RFID," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2018.

[20] G. Wang, H. Cai, C. Qian, J. Han, X. Li, H. Ding, and J. Zhao, "Towards replay-resilient rfid authentication," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[21] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2014.

[22] Impinj, "R420 readers," https://www.impinj.com/library, 2010.

[23] T. Foltýn, "Face unlock on many android smartphones falls for a photo," https://www.welivesecurity.com/2019/01/10/face-unlock-many-android-smartphones-falls-photo/, 2019.

[24] A. Greenberg, "Hackers say they've broken face id a week after iphone x release," https://www.wired.com/story/hackers-say-broke-face-id-security/, 2017.

[25] L. Tsang, J. A. Kong, and R. T. Shin, "Theory of microwave remote sensing," 1985.

[26] S. F. Kak and F. M. Mustafa, "Smart home management system based on face recognition index in real-time," in *International Conference on Advanced Science and Engineering*. IEEE, 2019, pp. 40–45.

[27] Q. Lin, L. Yang, C. Duan, and Y. Liu, "Revisiting reading rate with mobility: Rate-adaptive reading of COTS RFID systems," *IEEE Transactions on Mobile Computing (TMC)*, vol. 18, no. 7, pp. 1631–1646, 2019.

[28] C. Zhao, Z. Li, T. Liu, H. Ding, J. Han, W. Xi, and R. Gui, "Rf-mehndi: A fingertip profiled RF identifier," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2019.

[29] X. Lei and L. Sanglu, *Principle, Protocol and System Design of RFID*. Science Press, 2016.

[30] L. Yang, Y. Li, Q. Lin, H. Jia, X.-Y. Li, and Y. Liu, "Tagbeat: Sensing mechanical vibration period with cots rfid systems," *IEEE/ACM Transactions on Networking (TON)*, vol. 25, no. 6, pp. 3823–3835, 2017.

[31] M. Rizwan, "Lenet5a classic cnn architecture," https://www.datasciencecentral.com/lenet-5-a-classic-cnn-architecture/, 2018.

[32] SAMSUNG, "How does face recognition work on galaxy s20, s20+, s20 ultra, and z flip?" https://www.samsung.com/global/galaxy/what-is/face-recognition/.

[33] H. Chen, W. Wang, J. Zhang, and Q. Zhang, "Echoface: Acoustic sensor-based media attack detection for face authentication," *Internet of Things Journal (IoTJ)*, vol. 7, no. 3, pp. 2152–2159, 2019.

[34] Y. Chen, J. Sun, X. Jin, T. Li, R. Zhang, and Y. Zhang, "Your face your heart: Secure mobile face authentication with photoplethysmograms," in *Proceedings of the IEEE INTERNATIONAL Conference on Computer Communications (INFOCOM)*, 2017.

[35] Y. Cheng, X. Ji, J. Zhang, W. Xu, and Y. Chen, "Demicpu: Device fingerprinting with magnetic signals radiated by CPU," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2019.

[36] Z. Dehai, D. da, L. Jin, and L. Qing, "A pca-based face recognition method by applying fast fourier transform in preprocessing," in *Proceedings of the International Conference on Multimedia Technology (ICMT)*, 2013.

[37] C. Luo, Z. Yang, X. Feng, J. Zhang, H. Jia, J. Li, J. Wu, and W. Hu, "Rfaceid: Towards rfid-based facial recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 5, no. 4, pp. 1–21, 2021.

[38] B. Zhou, J. Lohokare, R. Gao, and F. Ye, "Echoprint: Two-factor authentication using acoustics and vision on smartphones," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[39] W. Xu, W. Song, J. Liu, Y. Liu, X. Cui, Y. Zheng, J. Han, X. Wang, and K. Ren, "Mask does not matter: anti-spoofing face authentication using mmwave without on-site registration," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2022.

[40] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, and T. Toft, "Privacy-preserving face recognition," in *International symposium on privacy enhancing technologies symposium*. Springer, 2009, pp. 235–253.

[41] A. Company, "About face id advanced technology," https://support.apple.com/en-us/HT208108, 2022.

[42] D. Oberhaus, "iphone x's face id can be fooled with a 3d-printed mask," https://www.vice.com/en/article/qv3n77/iphone-x-face-id-mask-spoof, 2017.

[43] A. Huang, D. Wang, R. Zhao, and Q. Zhang, "Au-id: Automatic user identification and authentication through the motions captured from sequential human activities using RFID," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 3, no. 2, pp. 48:1–48:26, 2019.

[44] J. Li, C. Wang, A. Li, D. Han, Y. Zhang, J. Zuo, R. Zhang, L. Xie, and Y. Zhang, "Rhythmic rfid authentication," *IEEE/ACM*

Transactions on Networking (TON), pp. 1–14, 2022.

[45] J. Liu, X. Zou, J. Han, F. Lin, and K. Ren, "BioDraw: Reliable multi-factor user authentication with one single finger swipe," in *Proceedings of the IEEE/ACM International Symposium on Quality of Service (IWQoS)*, 2020.

[46] J. Ning, L. Xie, C. Wang, Y. Bu, F. Xu, D.-W. Zhou, S. Lu, and B. Ye, "Rf-badge: Vital sign-based authentication via rfid tag array on badges," *IEEE Transactions on Mobile Computing (TMC)*, 2021.

[47] J. Liu, X. Zou, F. Lin, J. Han, X. Xu, and K. Ren, "Handkey: Leveraging multiple hand biometrics for attack-resilient user authentication using COTS RFID," in *Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2021.

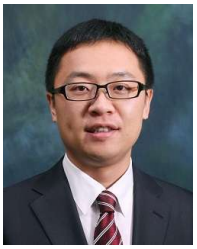
[48] C. Feng, J. Xiong, L. Chang, F. Wang, J. Wang, and D. Fang, "Rf-identity: Non-intrusive person identification based on commodity RFID devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 5, no. 1, pp. 9:1–9:23, 2021.

[49] J. Li, C. Wang, A. Li, D. Han, Y. Zhang, J. Zuo, R. Zhang, L. Xie, and Y. Zhang, "Rf-rhythm: Secure and usable two-factor RFID authentication," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, 2020.

Weiye Xu received her B.S. degree in computer science and technology from Wuhan University, China in 2019. She is now working toward the PhD degree at Zhejiang University. Her research interests include IoT security, wireless sensing and mobile computing. She is a student member of the IEEE.

Jianwei Liu received the BS degree from Northwestern Polytechnical University in 2018 and MS degree from Xi'an Jiaotong University in 2021. He is working toward the PhD degree at Zhejiang University. His research interests include RFID, mobile computing, and smart sensing. He is a student member of the IEEE.

Shimin Zhang is currently a master student majoring in Computer Science and Technology at Zhejiang University. Her research interests include smart sensing and wireless system security.



Yuanqing Zheng is currently an Associate Professor with the Department of Computing in Hong Kong Polytechnic University. He received the B.S. degree in Electrical Engineering and the M.E. degree in Communication and Information System from Beijing Normal University, Beijing, China, in 2007 and 2010 respectively. He received the PhD degree in School of Computer Engineering from Nanyang Technological University in 2014. His research interest includes Wireless Networking and Mobile Computing, Acoustic and RF Sensing, and Internet of Things (IoT). He is a member of IEEE and ACM.

Feng Lin received the Ph.D. degree from the Department of Electrical and Computer Engineering, Tennessee Technological University, USA, in 2015. He is currently a Professor with the School of Cyber Science and Technology, College of Computer Science and Technology, Zhejiang University, China. He was an Assistant Professor with the University of Colorado Denver, USA, a Research Scientist with the State University of New York (SUNY) at Buffalo, USA, and an Engineer with Alcatel-Lucent (currently, Nokia). His current research interests include mobile sensing, wireless sensing, Internet of Things security, biometrics, and AI security. Dr. Lin was a recipient of the ACM SIGSAC China Rising Star Award, the Best Paper Awards from ACM MobiSys'20, IEEE Globecom'19, IEEE BHI'17, the Best Demo Award from ACM HotMobile'18, and the Best Paper Award Nomination from Infocom'21. He serves as an editor for IEEE Network and IEEE Access.

Fu Xiao received the PhD degree in computer science and technology from the Nanjing University of Science and Technology, Nanjing, China, in 2007. He is currently a professor and a PhD supervisor with the School of Computer, Nanjing University of Posts and Telecommunications. He has authored papers in research related international conferences, including INFOCOM, ICC, and Mobihoc, the IEEE Journal on Selected Areas in Communications, IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, and ACM Transactions on Embedded Computing Systems. His research interest includes the Internet of Things.

Jinsong Han received his Ph.D. degree in computer science from Hong Kong University of Science and Technology in 2007. He is now a professor at the School of Cyber Science and Technology, Zhejiang University. He is a senior member of the ACM and IEEE. His research interests focus on IoT security, smart sensing, wireless and mobile computing.