# On Privacy Preserving Anonymization of Finger-selfies

Aakarsh Malhotra[1]*, Saheb Chhabra[1]*, Mayank Vatsa[2], Richa Singh[2]
[1]IIIT-Delhi, India; [2]IIT Jodhpur, India
[1]{aakarshm, sahebc}@iiitd.ac.in, [2]{mvatsa, richa}@iitj.ac.in

## Abstract

*With the availability of smartphone cameras, high speed internet, and connectivity to social media, users post content on the go including check-ins, text, and images. Privacy leaks due to posts related to check-ins and text is an issue in itself, however, this paper discusses the potential leak of one's biometric information via images posted on social media. While posting photos of themselves or highlighting miniature objects, users end up posting content that leads to an irreversible loss of biometric information such as ocular region, fingerprint, knuckle print, and ear print. In this paper, we discuss the effect of the loss of the finger-selfie details from social media. We demonstrate that this could potentially lead to matching finger-selfies with livescan fingerprints. Further, to prevent the leak of the finger-selfie details, we propose privacy preserving adversarial learning algorithm. The algorithm learns a perturbation to prevent the misuse of finger-selfie towards recognition, yet keeping the visual quality intact to highlight the minuscule object. The experiments are presented on the ISPFDv1 database. Further, we propose a new publicly available Social-Media Posted Finger-selfie (SMPF) Database, containing 1,000 finger-selfie images posted on Instagram.*

## 1. Introduction

With the advent of technology, the internet, and smart devices are easily accessible. These two factors facilitate users to post content on social media easily. Users utilize these posted content for varied tasks such as sharing their experiences, thoughts, discussions, blogging, and connecting with people of common interests. These tasks can be achieved via sharing of texts, images, voice notes, and check-ins. Of these, images are one of the most popular ones. Users utilize photos to post content such as selfies, profile pictures, food, and travel blogging. During this process, they post many photos of themselves intentionally or unintentionally. As shown in Figure 1, an image of a user

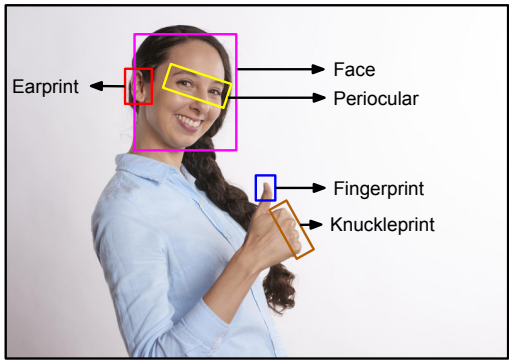


Figure 1. Potential biometric modalities that could be unintentionally revealed by social media users on uploading their photos. Image taken from the Internet; Link: https://tinyurl.com/yazwkwuf

can reveal many biometric modalities such as face, fingerprint, and ocular which are actively used for authentication applications. Therefore, these social media posts may become a source of irreversible privacy leak with respect to biometric information.

Images of finger can be captured via phones (known as finger-selfie[1]) and similar to fingerprints, the ridge-valley patterns present in finger-selfies can also be used for recognition. As shown in Figure 2, users may expose the ridge-valley details of the fingerprint while posting images of miniature objects on their hands. Recent studies have showed that unintended access to ridge-valley patterns could lead to unauthorized access, phone unlocking, and bank frauds via the use of silicone molds. Hern [11] showed how fingerprints of a German minister could be reconstructed using DSLR acquired thumbs-up image from a distance of 3 meters. Recently, a suspect was convicted based on a WhatsApp circulated image of him holding drugs [27]. Similarly, access to ridge-valley details from latent impression of the surface of smartphones [6, 15] or other sources can be used to generate 3D silicone fingerprints [20] (or presentation attack) to spoof the recognition system [21, 26]. Hence, it is essential to anonymize ridge-valley details while the picture is posted on social media. In this paper, we therefore focus on privacy preserving

*Equal contribution by the student authors.

[1]Finger-selfie: An image of frontal region of the finger.

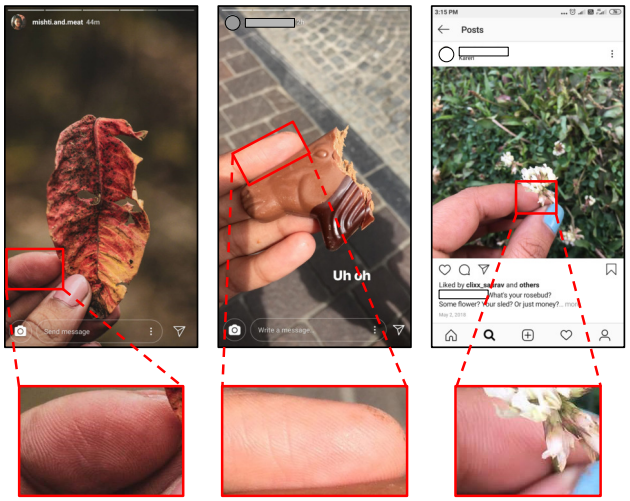


Figure 2. Social media users implicitly revealing their fingerprint details while posting image of miniature objects. Photos shown after taking consent from users. Credits for image one with dry leaf: Shramona Poddar.

anonymization of finger-selfies.

In the literature, researchers have proposed several algorithms [2, 7, 8, 12, 18] to anonymize private and sensitive data. Majority of these algorithms are based on the concept of $k-$anonymity [23] where the attributes of an individual cannot be distinguished from at least $k-1$ other individuals. Howbeit, there are several limitations of $k$-anonymity based algorithms including loss of data utility while preserving privacy. For social media applications, achieving both privacy and data utility is a challenging task. In order to address this problem, a privacy preserving algorithm is proposed which anonymizes the input data while preserving the data utility. For social media applications, the data utility corresponds to visual appearance of the image.

For fingerprints, limited attention has been provided towards adversarial perturbations. To the best of our knowledge, this is the first study that focuses on anonymizing finger-selfies or has aimed to preserve the exposed ridge-valley details over the images uploaded on the social media. Grosz et al. [10] studied the vulnerability of each module of fingerprint recognition towards minutia perturbations. Their algorithm introduced perturbations on the minutia feature set and not the fingerprint image itself. Fernandes et al. [5] also perturbed livescan fingerprints to fool a VGG19 based system. Marrone and Sansone [17] introduced adversarial perturbation in fingerprints. While recognition performance did reduce, the perturbation profoundly affected the visual quality of the fingerprint images.

Deep learning models are still considered as gray or black-box models. These models learn complex features from an image for recognition or classification. For (contactless) fingerprint recognition using deep learning models, there is no such finding that deep learning models learn only ridge-valley patterns. As highlighted by Kumar and Zhou [13], identifiable details can be found from all three finger segments. Further, solely blurring or pixelization may remain ineffective since the de-identified probe can still be matched with a de-identified gallery [9]. It is important to anonymize identity in all finger-skin regions. It can be achieved when we target perturbation only in skin regions, without altering other object in the image. In this paper, we propose a privacy preserving algorithm for finger-selfies that can be applied as a pre-processing step during social media posting. The algorithm operates directly on images and preserves visual appearance. Since the social media platform would lack the corresponding ground truth template, the proposed algorithm works only using the impostor pairs. The proposed anonymization algorithm affects smaller number of pixels as it uses skin-color for localized anonymization. Finally, we also present an UnShared Siamese model to match finger-selfies with fingerprints. In order to conduct the real world evaluation, we have also collected and release a database of 1,000 finger-selfies from social media. Each of these finger-selfies consists of visible ridge-valley details of the finger. While the proposed anonymization algorithm is shown for finger-selfies as a case study, the proposed algorithm is generalized and may also be used for face or knuckle-print.

## 2. Privacy Preserving Anonymization

In this research, we have proposed a privacy preserving algorithm based on the concept of adversarial perturbation. Let $x_j$ be an original finger-selfie of subject $j$ to be anonymized. The pixels of image $x_j$ are in the range of 0 to 1. The task is to learn the perturbation $p_j$ such that on adding $p_j$ to finger-selfie $x_j$, the identity is anonymized while preserving the overall visual quality. Mathematically, it is written as:

$$a_j = x_j + p_j \tag{1}$$

where, $a_j$ is the output anonymized image. As mentioned above, $x_j$ is in the range of 0 to 1, therefore, output anonymized image $a_j$ should also be in the range of 0 to 1. For this purpose, the following transformation function is applied.

$$a_j = \frac{1}{2}(tanh(x_j + p_j) + 1) \tag{2}$$

In this research, we assume that siamese network is used for fingerprint recognition. Let $g(.)$ be the siamese network used for fingerprint recognition which takes a pair of image as input and outputs whether the pair belongs to same identity or not. Mathematically, it is written as:

$$\mathcal{S}_{il,js} = g(x_{il}, x_{js}) \tag{3}$$

where, $x_{il}$ is the livescan fingerprint pertaining to the subject $i$ and $x_{js}$ is the finger-selfie image belonging to the

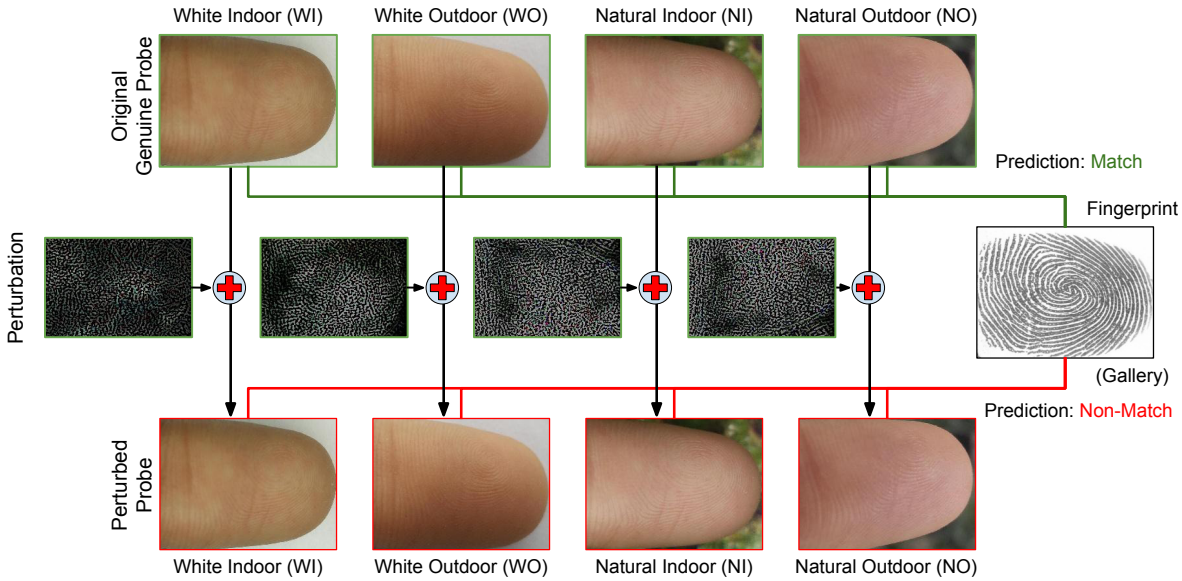


Figure 3. Original genuine finger-selfie images getting misclasified as non-match after adding the perturbation image. As seen in perturbations, the highest perturbation is learnt inside finger region and specifically near singular points. Noise amplified for illustration.

subject $j$. $S_{il,js}$ is the output matching score in the range of 0 to 1. The score 1 indicates the complete match while score 0 indicates the non-match of images $x_{il}$ and $x_{js}$. In order to anonymize the input image $x_{js}$ corresponding to siamese network $g(.)$, two scenarios are possible. In the first scenario, the input image with subject $j$ is anonymized corresponding to other images of the same subject $j$. In the second scenario, the input image is anonymized using images of subjects other than the subject $j$. In a real world scenario, the images of same subject or genuine pairs are very limited and might not be available openly. Yet, the images of other subjects or impostor pairs are easily available. Therefore, in this research, we have anonymized the image corresponding to the second scenario. The details of the optimization are discussed below.

## 2.1. Optimization

To anonymize the input image $x_{js}$ corresponding to siamese network $g(\cdot)$, random impostor pairs are generated with image $x_{js}$ of subject $j$ and other images $x_{il}$ of many different subjects $i$, where $i \neq j$. Let $m$ be the number of impostor pairs generated to anonymize image $x_{js}$. Therefore, to anonymize $x_{js}$, the distance between the output impostor pairs score $S_{il,js}$ and the target score $c = \frac{1}{2}$ is minimized. Mathematically, it is written as:

$$\text{minimize} \sum_{i=1}^{j=m} D\left(c, \mathcal{S}_{il,js}\right) \quad i \neq j \tag{4}$$

where $D(\cdot)$ is the distance metric. The above equation enforces the impostor matching scores towards $c = \frac{1}{2}$. The reason for choosing $c = \frac{1}{2}$ is to shift the impostor distribution towards genuine score distribution and to introduce multiple traits of different non-match identities into the finger-selfie image. This would implicitly reduce the match criterion of finger-selfie $x_{js}$ when it is matched with its true pair $x_{il}$ $(i = j)$. Further, enforcing the impostor matching scores towards $\frac{1}{2}$ would not add any dominating features of any one particular identity. Cases with higher dissimilarity (scores towards 0) would imply more artifacts are added to finger-selfie, thus degrading the performance even further. In this research, the following function is used to minimize the distance between the output score $\mathcal{S}_{il,js}$ and constant $c$.

$$\text{minimize} \sum_{i=1}^{j=m} max(0, c - \mathcal{S}_{il,js}) \tag{5}$$

The function used in above equation enforce only the correctly classified impostor pair score $S_{il,js}$ towards $c = \frac{1}{2}$. If the imposter pair gets misclassified, the output score will be greater than $c = \frac{1}{2}$. Thus, the above equation will not be optimized for the same. Additionally, for two different finger-selfies of the same subject, each image is optimized corresponding to different random impostor pairs. This results in different learned perturbations and artifacts for both the images. Therefore, both the images have different identity information and do not match with each other.

In social media applications, preserving the visual appearance of an anonymized image $x_{js}$ is also important. Inspired from Carlini and Wagner [1], the $l_2$ distance between the input image $x_{js}$ and output anonymized image $a_{js}$ obtained by adding $p_{js}$ is minimized. Minimizing $l_2$ distance

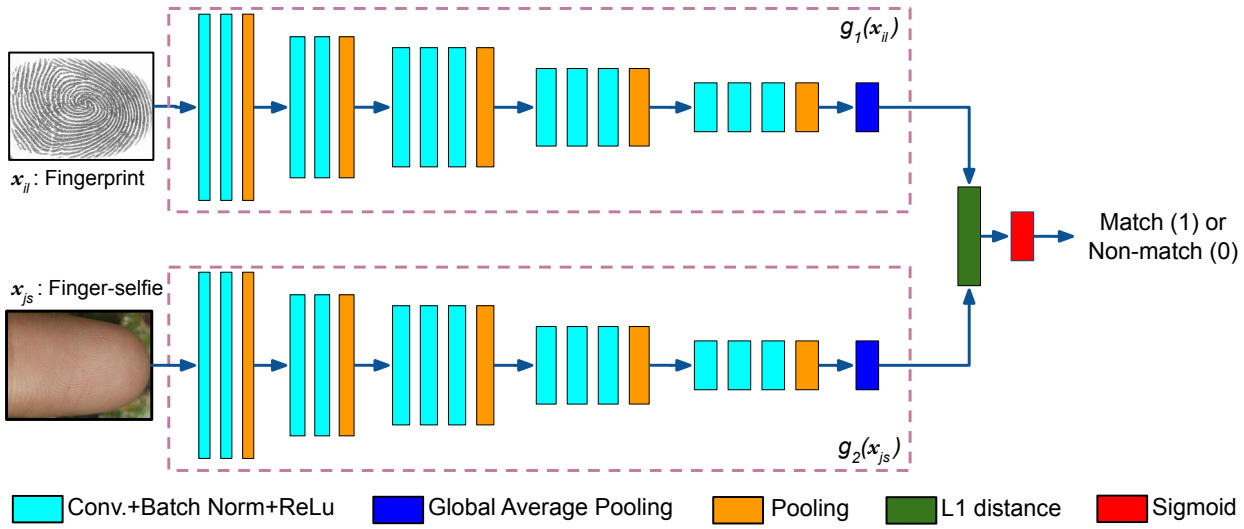


Figure 4. Matching finger-selfies against a gallery of livescan fingerprint using UnShared weights architecture. (Best viewed in color).

while fooling classifiers results in output perturbed images which are visually indistinguishable from the original image [1]. Further, Euclidean distance is known to be sensitive to outliers and therefore, minimizing euclidean distance produces smoothed output. Mathematically, it is written as:

$$\text{minimize } ||\mathbf{x}_{js} - \mathbf{a}_{js}||_2^2 \tag{6}$$

On combining Equation 5 and 6, the final optimization function is written as:

$$\text{minimize } \sum_{i=1}^{j=m} max(0, c - \mathcal{S}_{il,js}) + ||x_{js} - a_{js}||_2^2 \tag{7}$$

It is important to note that the above function is optimized corresponding to the variable $p_{js}$.

### 2.2. Skin Color based Optimization

In many cases, it is observed that the background region is present along with the finger region in the finger-selfie. The presence of background region can hinder the performance and results in learned perturbation focusing on background region instead of finger region. In order to resolve this problem, we have generated a mask $m_{js}$ using skin color segmentation corresponding to finger region in the finger-selfie. The mask $m_{js}$ is applied on the perturbation $p_{js}$ while optimization. Mathematically, it is written as:

$$a_j = \frac{1}{2}(tanh(x_j + mp_j) + 1) \tag{8}$$

where, the mask $m_{js}$is is multiplied with perturbation $p_{js}$ element-wise. Using the above Equation in place of Equation 2 during optimization, the learned perturbation focusing on finger region only.

## 3. Finger-selfie Recognition Using UnShared Siamese Model

This section proposes a Siamese-like framework to match finger-selfies to livescan fingerprints. Siamese CNNs [3] have shown excellent performance on image [14] and biometric recognition [22]. It has two CNNs sharing the same network structure and weights, and input consists of match and non-match image pairs. Based on such architecture, Lin and Kumar [16] showed a multi-view Siamese CNN for matching 3D fingerprints. As seen in Figure 4, visually, the inputs (fingerprint vs finger-selfie) differ significantly, resulting in different distributions and statistics in the input space. Consequently, processing different domains of input from the same weights to obtain discriminatory features would be unfair.

As shown in Figure 4, we propose UnShared Siamese Architecture with two streams of VGG16 architecture. The idea is inspired from DeepFace, where the locally connected layer learns different set of filters for every location in the feature map [25]. Unlike the traditional siamese framework, the two streams have disjoint weights and are optimized separately in the UnShared Siamese Network. Once the training is completed, the two streams have a different set of weights as a transformation function of input: $g_1(X_1)$ and $g_2(X_2)$. Fine-tuning different siamese streams allow the network to learn appropriate weights and extract relevant discriminatory features. Since VGG16 has shown its effectiveness in extracting discriminatory information from finger-selfies [4], it is selected as a base network for each stream. Each of these streams is pre-trained on ImageNet dataset classification weights.

We assume that each of the sub-networks keeps the discriminatory features intact in its feature representation. As

mentioned in Section 2, for any input pair to the network, $x_{il}$ image represents the livescan fingerprint from $i^{th}$ person and $x_{js}$ represents the finger-selfie image from $j^{th}$ person. Intuitively, if pairs $\{x_{il}, x_{js}\}$ are from same identity ($i = j$), the representation $g_1(x_{il})$ and $g_2(x_{js})$ should be close to each other in feature space. In other words, the distance between $g_1(x_{il})$ and $g_2(x_{js})$ should be low. Similarly, for pairs arising from different identities ($i \neq j$), distance between $g_1(x_{il})$ and $g_2(x_{js})$ should be high. The model calculates the $l_1$ distance between representations $g_1(x_{il})$ and $g_2(x_{js})$ as shown below:

$$d_{x_{il},x_{js}} = \|g_1(x_{il}) - g_2(x_{js})\|_1^1 \tag{9}$$

A sigmoid unit classifies the distance $d_{x_{il},x_{js}}$ as match (ideal score $\mathcal{S}_{il,js} = 1$; for $i = j$) or non-match (ideal score $\mathcal{S}_{il,js} = 0$; for $i \neq j$). For testing, the input for the trained model is a fingerprint-finger-selfie pair. The model evaluates if these two images belong to the same identity or not.

# 4. Experimental Details

## 4.1. Database and Protocol

To evaluate the effectiveness of the proposed concept, we utilize IIITD SmartPhone Fingerphoto Database v1 (ISPFDv1) [19]. This database has fingerphotos[2] acquired under different scenarios of (i) White Indoor (WI), (ii) White Outdoor (WO), (iii) Natural Indoor (NI), and (iv) Natural Outdoor (NO). These scenarios cover potential cases of where the fingerphoto can be captured. Further, the ISPFDv1 also contains the corresponding livescan fingerprints of the users. We use the fingerprint-fingerselfie pair to train a model for recognizing finger-selfie ridge details against the ridge details present in livescan fingerprints. In a subject disjoint manner, pairs from 50% subjects of the ISPFDv1 are used to train the recognition model. We then report the original recognition performance on fingerprint-fingerselfie pairs from the remaining 50% subjects of the testing set. The finger-selfie from the testing pairs is then perturbed using the adversarial model to hamper the recognition performance.

At the time of writing this paper, searching for hashtag *"#miniature"*, *"#finger"*, and *"#myhand"* reveals 4,344,151, 1,254,070, and 212,027 Instagram posts respectively. These posts include people showing their wedding rings, highlighting miniature objects such as a dandelion flower, showing nail paints and tattoos. Most of these posts contain single or multiple fingers in the uploaded photograph. Few of these photographs also contain clearly visible ridge-valley details. Using these three hashtags, we collect **Social-Media Posted Finger-selfie (SMPF) database**.

[2]Fingerphoto and finger-selfie in our context are same as they are captured by the smartphone camera by the user themselves.

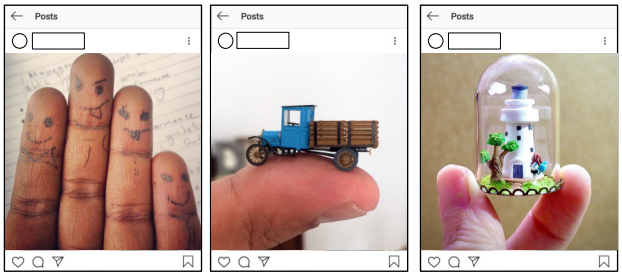

Figure 5. Sample images from the proposed SMPF Database.

The database consists of 1,000 images, each of which is collected from public Instagram posts. These images are handpicked so that ridge-valley details in proximal phalanx is visible. As a part of this research, we would also share links to these Instagram posts[3]. Few examples of such images can be seen in Figure 5. Since these images are taken from social media, we do not have corresponding livescan fingerprints. Nevertheless, since the proposed adversarial algorithm relies on impostor pairs, we create impostor pairs by fetching livescan fingerprints from ISPFDv1. Thus, the recognition model trained on ISPFDv1 is used for obtaining impostor score distribution for pairs with ISPFDv1 fingerprint and SMPF finger-selfie images.

## 4.2. Implementation Details

The details for training finger-selfie-to-fingerprint matching algorithm and perturbation learning are discussed below. Both models are trained on Nvidia GeForce RTX 2080Ti and implemented in Keras. For each of the models, the input is a fingerprint-finger-selfie image pair with resolution 270×200×3. The finger-selfie is compressed thrice with different compression rates. The lower image resolution and varied compression rates closely resemble the quality of finger regions in the pictures uploaded on social media.

**Perturbation Learning:** To learn the perturbation, learning rate is set to 0.001 and the anonymization is performed corresponding to five impostor pairs. The number of iterations used for anonymizing each image is 20.

**Finger-selfie-to-fingerprint Matching Model:** To train the matching model, each stream of the UnShared Siamese Architecture is initialized using a pre-trained VGG16 model. The network is optimized over binary cross-entropy loss for 60 epochs with a batch size of 16. The model uses Adam optimizer with a learning rate of $5 \times 10^{-5}$. The network is trained for coarsely cropped finger-selfie matching.

# 5. Results

In order to evaluate the performance of the proposed algorithm, anonymization is performed for four different cases. They are described as follows:

[3]The file containing links of the posts can be downloaded from: http://iab-rubric.org/resources/smpf.html

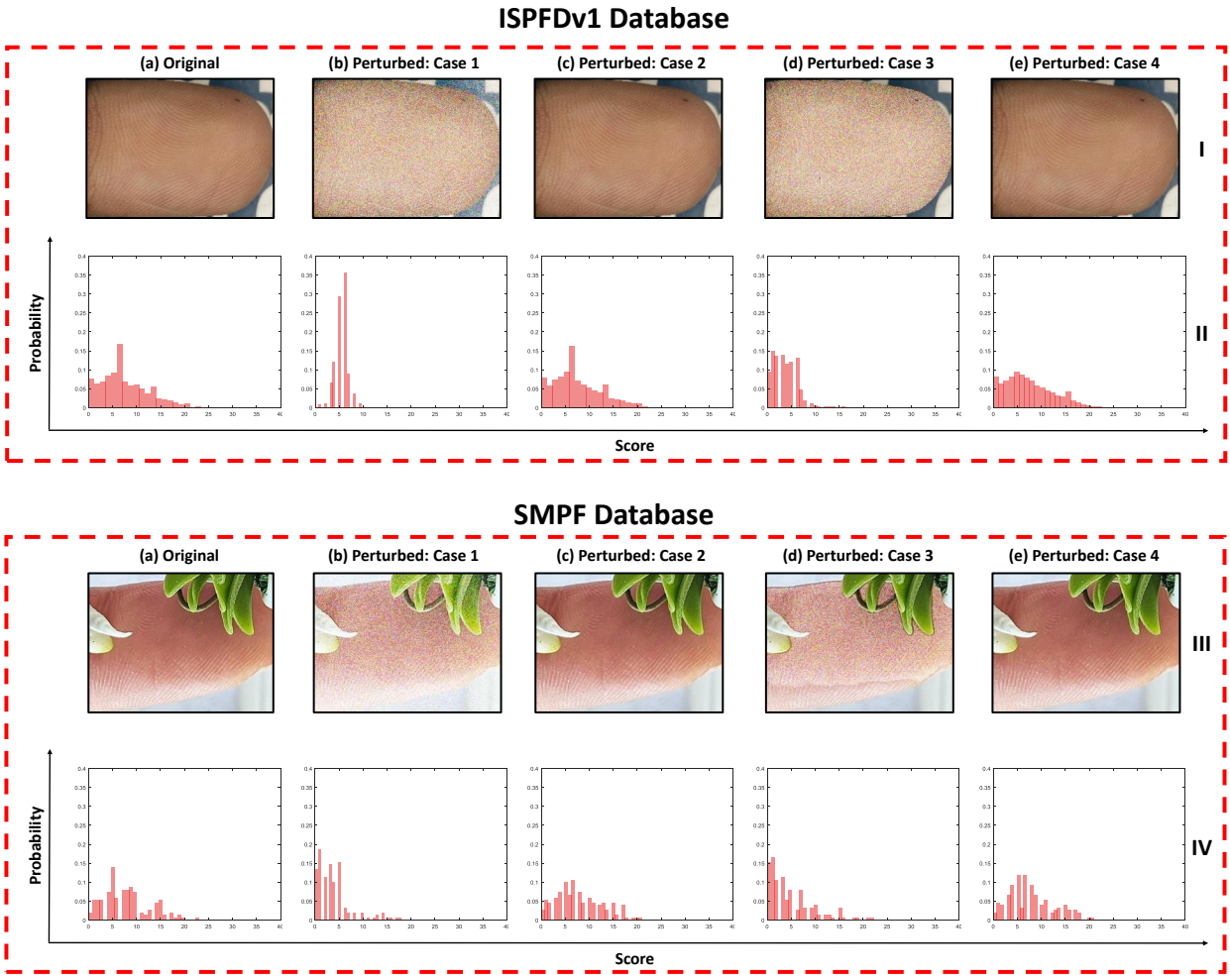


Figure 6. Effect of perturbation on finger-selfie quality illustrated visually (row I and row III) and with NFIQ 2.0 scores (row II and row IV) under different cases. (a) Original finger-selfie, (b) Case 1: Anonymization of the whole image without preserving visual appearance, (c) Case 2: Anonymization of the whole image while preserving visual appearance, (d) Case 3: Anonymizing finger-region only without preserving visual appearance, and (e) Case 4: Anonymizing finger-region only while preserving visual appearance. Images from ISPFDv1 [19] and the proposed SMPF database.

- **Case 1:** We anonymize the whole image including finger and background regions. In this case, the visual appearance of the image is not preserved.
- **Case 2:** Anonymization is performed for the whole image while preserving the visual appearance of the image.
- **Case 3:** It deals with anonymization of finger region only in the finger-selfie and the visual appearance is not preserved.
- **Case 4:** Similar to Case 3, Case 4 deals with finger region only while preserving the visual appearance.

### 5.1. Image Quality

The effect of anonymization by adding perturbation can be seen in Figure 6 (I). As seen from the perturbed Case 2 and Case 4 images, the constraint to preserve visual quality ensures that the details remain preserved visually. This results in reduced recognition performance, as discussed in next subsection. Further, to ensure that minutiae details get distorted with perturbation, Figure 6 (b) and Figure 6 (d) show samples where we allow distortion to happen in finger-selfie. Furthermore, with a use case of localized privacy preservation, perturbation is added only in skin regions in Cases 3 and 4. Similar results can be visually validated from Figure 6 (III) for the proposed SMPF database. The proposed method can be extended to other skin-based biometrics such as face, ear, and knuckle print.

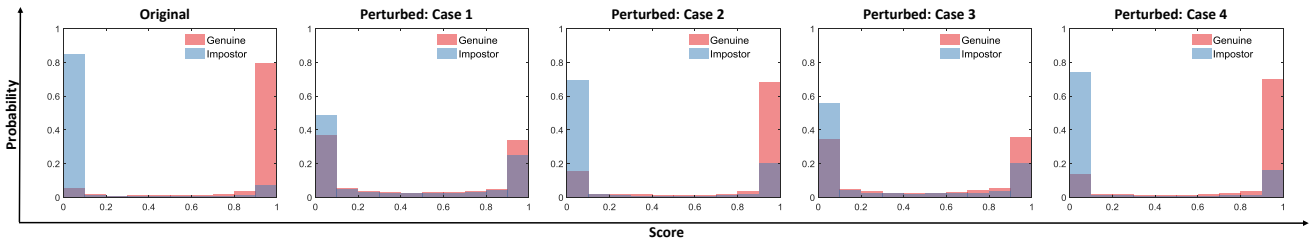


Figure 7. Case-wise matching score distribution for original and perturbed finger-selfie for ISPFDv1 database (Best viewed in color).

Table 1. Effect of adding perturbation on different metrics under different use case scenarios. Results are reported on the ISPFDv1 database.

| Metric | Original | Perturbed | | | |
|---|---|---|---|---|---|
| | | Case 1 | Case 2 | Case 3 | Case 4 |
| Model performance (%) | 89.09 | 55.47 | 75.76 | 59.69 | 79.54 |
| EER (%) | 10.89 | 44.24 | 24.47 | 39.32 | 20.45 |
| Genuine Score (Mean) | 0.88 | 0.49 | 0.76 | 0.51 | 0.78 |
| Impostor score (Mean) | 0.11 | 0.38 | 0.26 | 0.32 | 0.21 |
| SSIM (Mean) | - | 0.24 | 0.99 | 0.41 | 0.99 |
| Perturbation Magnitude | - | 0.1370 | 0.0031 | 0.1050 | 0.0030 |
| NFIQ 2.0 quality score (Mean) | 7.67 | 5.39 | 7.67 | 3.66 | 7.60 |

To analyze quality based on fingerprint specific traits, we utilize NFIQ 2.0 metric [24]. Anonymizing finger-selfie under different cases has different effects on fingerprint specific quality. As seen in Table 1, the original mean quality score for the original finger-selfie is 7.67. From Figure 6 (II.a), we observe that the quality scores range till 30. It is our assertion that the lower scores arise due to intentionally degraded quality by multiple JPEG compression and lower resolution. After adding perturbation, when we constrain the learning to keep visual details intact (Cases 2 and 4), we observe that fingerprint image quality remains the same. However, on removing visual constraints (Cases 1 and 3), the NFIQ quality score degrades. The same can be validated from score distributions in Figure 6 (II.d). In scenarios where we want traditional fingerprint minutiae extractor to fail, we may utilize Case 3. Case 3 would adversely affect ridge valley details only in skin areas. However, Cases 2 and 4 aims to only prevent recognition against deep models while preserving the visual quality of finger-region.

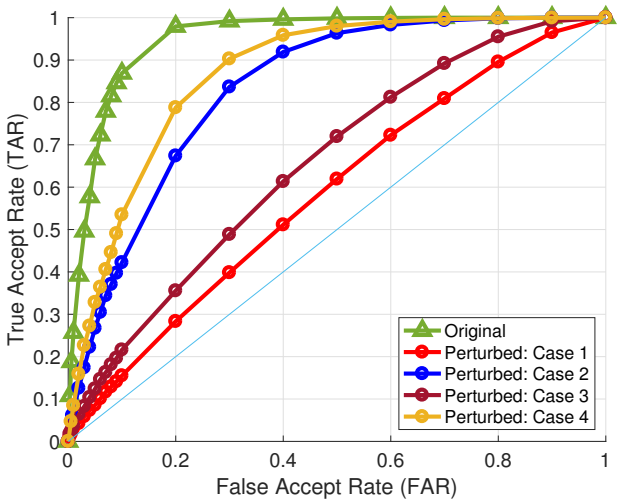


Figure 8. ROC curves for finger-selfie recognition using the proposed recognition model, before and after adding perturbation to fingerphotos. (Best viewed in color).

## 5.2. Recognition

In this subsection, we study the performance drop after the privacy preserving anonymization. From Table 1, it can be observed that the proposed UnShared Siamese Model provides a performance of 89.09% on the ISPFDv1 Database. The ROC curves, before and after adding perturbation, is shown in Figure 8.

On adding perturbation under Case 1 and Case 3 scenario, we allow the network to deteriorate visual details of the image. From Figure 6 (I and III), we observe that ridge-valley details are barely visible in Case 1 and 3. With the lack of discriminative details, the model fails to recognize pairs correctly and yields an accuracy of 55.47% and 59.69%, respectively. The same can be validated by the lower peaks of genuine and impostor score in Figure 7.

Further, in Case 2 and Case 4, we constraint the perturbation to preserve data utility, i.e., while anonymizing, the adversarial network preserves the visual details. From Figure 6 (I and III), we observe that ridge-valley details are intact in Case 2 and 4. However, small perturbations from

Table 2. Effect of adding perturbation on different metrics under different use case scenarios. Results are reported on the SMPF database.

| Metric | Original | Perturbed | | | |
|---|---|---|---|---|---|
| | | Case 1 | Case 2 | Case 3 | Case 4 |
| Model performance (%) | 80.00 | 58.70 | 56.70 | 63.30 | 64.00 |
| Impostor score (Mean) | 0.20 | 0.42 | 0.43 | 0.37 | 0.38 |
| SSIM (Mean) | - | 0.35 | 0.99 | 0.55 | 0.99 |
| Perturbation Magnitude | - | 0.1178 | 0.0031 | 0.0895 | 0.0031 |
| NFIQ 2.0 quality score (Mean) | 7.78 | 3.77 | 7.76 | 4.73 | 7.78 |

multiple impostor classes ensure that the deep model classifies the samples incorrectly. Accordingly, the model performance reduces to 75.76% in Case 2 and 79.54% in Case 4. The lesser reduction in performance in Case 4 can be attributed to False Negatives in the skin color map. Due to some skin regions getting classified as non-skin, the adversarial algorithm does not alter particular ridge-valley details. These details help the UnShared Siamese model to classify the samples correctly.

Due to the lack of fingerprints for the SMPF database, we create pairs for SMPF with fingerprints from ISPFDv1. Since only impostor pairs can be created, we lack genuine scores due to which the equal error rate cannot be calculated. Nevertheless, prediction scores below 0.5 can be classified as True Negative and above 0.5 can be classified as False Positive. The rest of the metrics for the original and perturbed SMPF images are also shown in Table 2. We can observe that the impostor score shifts towards 0.5 after adding perturbation. Consequently, a fall in recognition performance for the actual classes can also be expected.

## 6. Conclusion and Future Work

This paper focuses on fingerselfie anonymization for preserving the privacy of individuals before posting images on the social media. We show how a finger-selfie shared on social media could potentially be matched with livescan fingerprints using the proposed UnShared VGG16 Siamese Architecture. To prevent the misuse of finger-selfie towards unauthorized recognition, we propose an adversarial learning-based perturbation algorithm. Multiple different experiments are performed that demonstrate that the learned perturbation can either be added in the complete image or locally into the skin regions, which leads the fingerselfie to be incorrectly classified. The experiments performed on both ISPFDv1 and the proposed Social-Media Posted Finger-selfie (SMPF) database demonstrate the effectiveness of the algorithms. Further analysis on selecting impostor samples, perturbation magnitude, and robustness towards fingerprint enhancement techniques could pave the path for future work. The concept of secure and anonymized posting of social media images can be further extended to other biometric modalities.

## 7. Acknowledgment

Aakarsh Malhotra is partially supported by Visvesvaraya Ph.D. Fellowship. Mayank Vatsa is supported through the Swarnajayanti Fellowship by the Government of India. Authors extend their gratitude towards Alka Malhotra for her assistance in dataset collection.

## References

[1] Nicholas Carlini and David Wagner. Towards evaluating the robustness of neural networks. In *IEEE Symposium on Security and Privacy*, pages 39–57, 2017.

[2] Saheb Chhabra, Richa Singh, Mayank Vatsa, and Gaurav Gupta. Anonymizing k facial attributes via adversarial perturbations. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pages 656–662, 7 2018.

[3] Sumit Chopra, Raia Hadsell, and Yann LeCun. Learning a similarity metric discriminatively, with application to face verification. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 539–546, 2005.

[4] Shaan Chopra, Aakarsh Malhotra, Mayank Vatsa, and Richa Singh. Unconstrained fingerphoto database. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 517–525, 2018.

[5] Steven Fernandes, Sunny Raj, Eddy Ortiz, Iustina Vintila, and Sumit Kumar Jha. Directed adversarial attacks on fingerprints using attributions. In *IEEE International Conference on Biometrics*, pages 1–8, 2019.

[6] Ines Goicoechea-Telleria, Ana Garcia-Peral, Anas Husseis, and Raul Sanchez-Reillo. Presentation attack detection evaluation on mobile devices: Simplest approach for capturing and lifting a latent fingerprint. In *IEEE International Carnahan Conference on Security Technology*, pages 1–5, 2018.

[7] Ralph Gross, Edoardo Airoldi, Bradley Malin, and Latanya Sweeney. Integrating utility into face de-identification. In *International Workshop on Privacy Enhancing Technologies*, pages 227–242. Springer, 2005.

[8] Ralph Gross, Latanya Sweeney, Jeffrey Cohn, Fernando De la Torre, and Simon Baker. Face de-identification. In *Protecting privacy in video surveillance*, pages 129–146. Springer, 2009.

[9] Ralph Gross, Latanya Sweeney, Fernando De la Torre, and Simon Baker. Model-based face de-identification. In *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pages 161–161, 2006.

[10] Steven A Grosz, JJ Engelsma, Nicholas G Paulter, and Anil K Jain. White-box evaluation of fingerprint matchers: Robustness to minutiae perturbations.

[11] Alex Hern. Hacker fakes German minister's fingerprints using photos of her hands. https://tinyurl.com/y5qsrwlu. Accessed: 08-Nov-2019.

[12] Amin Jourabloo, Xi Yin, and Xiaoming Liu. Attribute preserved face de-identification. In *IEEE International Conference on Biometrics*, pages 278–285, 2015.

[13] Ajay Kumar and Yingbo Zhou. Contactless fingerprint identification using level zero features. In *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pages 114–119, 2011.

[14] BG Kumar, Gustavo Carneiro, Ian Reid, et al. Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5385–5394, 2016.

[15] Hoyeon Lee, Seungyeon Kim, and Taekyoung Kwon. Here is your fingerprint! Actual risk versus user perception of latent fingerprints and smudges remaining on smartphones. In *Annual Computer Security Applications Conference*, pages 512–527, 2017.

[16] Chenhao Lin and Ajay Kumar. Contactless and partial 3d fingerprint recognition using multi-view deep representation. *Pattern Recognition*, 83:314–327, 2018.

[17] Stefano Marrone and Carlo Sansone. Adversarial perturbations against fingerprint based authentication systems. In *IEEE International Conference on Biometrics*, pages 1–6, 2019.

[18] Elaine M Newton, Latanya Sweeney, and Bradley Malin. Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering*, 17(2):232–243, 2005.

[19] Anush Sankaran, Aakarsh Malhotra, Apoorva Mittal, Mayank Vatsa, and Richa Singh. On Smartphone Camera based Fingerphoto Authentication. In *IEEE International Conference on Biometrics Theory, Applications and Systems*, pages 1–7, 2015.

[20] Jan Spurnỳ, Michal Doleel, Ondrej Kanich, Martin Drahanskỳ, and Koichi Shinoda. New materials for spoofing touch-based fingerprint scanners. In *IEEE International Conference on Computer Application Technologies*, pages 207–211, 2015.

[21] Chris Stein, Vincent Bouatou, and Christoph Busch. Video-based fingerphoto recognition with anti-spoofing techniques with smartphone cameras. In *IEEE International Conference of the Biometrics Special Interest Group*, pages 1–12, 2013.

[22] Yi Sun, Yuheng Chen, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation by joint identification-verification. In *Advances in Neural Information Processing Systems*, pages 1988–1996, 2014.

[23] Latanya Sweeney. k-anonymity: A model for protecting privacy. *IJUFKS*, 10(05):557–570, 2002.

[24] Elham Tabassi. Development of NFIQ 2.0. https://www.nist.gov/services-resources/software/development-nfiq-20. Accessed: 30-Mar-2020.

[25] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014.

[26] Archit Taneja, Aakriti Tayal, Aakarsh Malhorta, Anush Sankaran, Mayank Vatsa, and Rieha Singh. Fingerphoto spoofing in mobile devices: a preliminary study. In *IEEE International Conference on Biometrics Theory, Applications and Systems*, pages 1–7, 2016.

[27] Chris Wood. WhatsApp drug dealer caught by 'groundbreaking' work. http://www.bbc.com/news/uk-wales-43711477. Accessed: 11-Nov-2019.