# FaceIDP: Face-IDentification Privacy under the Compressed Sensing Framework

Lu Ou, *Member, IEEE,* Shaolin Liao*, *Senior Member, IEEE*, Zheng Qin, *Member, IEEE,* Yuan Hong, *Senior Member, IEEE*, and Dafang Zhang

***Abstract*—In FaceID era, large number of facial images could be used to breach the FaceID system, which demands effective FaceID privacy protection of the facial images for widespread adoption of FaceID technique. In this paper, to our best knowledge, we take the first step to systematically study such important FaceID privacy issue, under the framework of Compressed Sensing (CS) for fast facial image transmission. Specifically, we develop the Face-IDentification Privacy (FaceIDP) approach to protect the facial images from being used by the adversary to breach some FaceID system. First, a Dictionary Learning neural Network (DLNet) has been developed and trained with facial images database, to learn the common dictionary basis of the facial image database. Then, the encoding coefficients of the facial images are obtained. After that, the sanitizing noise is added to the encoding coefficients, which obfuscates the FaceID feature vector that is used to identify the FaceID. We have also proved that the FaceIDP is $\varepsilon$-differentially private. More importantly, optimal noise scale parameters have been obtained via the Lagrange Multiplier (LM) method to achieve better data utility for a given privacy budget $\varepsilon$. Finally, substantial experiments have been conducted to validate the efficiency of the FaceIDP with two real-life facial image databases, *i.e.*, the LFW (Labeled Faces in the Wild) database and the PubFig database, and the results show that it outperforms other commonly used Differential Privacy (DP) approaches.**

***Index Terms*—FaceID, Face Identification Privacy (FaceIDP), Differential Privacy (DP), Compressed Sensing (CS), Dictionary Learning Neural Network (DLNet).**

## I. Introduction

FACE recognition has been extensively used as a biometric authentication method in many fields such as public safety, finance, e-commerce, *etc.*, due to its super convenience [1]. Also, in the 5G and beyond era where images and videos on the internet clouds can be transmitted and shared in real time and faster speed than ever [2], [3]. This poses great threat to the FaceID identification systems since the adversaries could combine a user's multiple facial images to form the 3D feature point cloud and breaches the FaceID system to identify the user of interest. This is especially true in the era of Artificial Intelligence (AI): through training large number of facial images of the users, face feature vectors could be learned accurately; then the face identification of the user is carried out through deep learning, leading to privacy leakage from mining information of the publicly shared facial images [4], [5], [6], [7]. Thus the FaceID systems face the real risk of being breached and FaceID is forbidden in many cities, such as San Francisco and Boston, in USA. Therefore, effective FaceID privacy protection technique is urgent for widespread adoption of FaceID applications.

However, there is lack of research on the privacy problem of FaceID protection from falsified 3D feature cloud points from machine learning of publicly released facial images. However, research on other privacy problems other than the FaceID privacy of publicly shared facial images exist. For example, in order to protect the facial image privacy, image obfuscation [8], [9], [10] such as pixelization and blurring, is adopted to protect image features. Unfortunately, these methods could be re-identified. To fix this problem, a differentially private pixelization is proposed [11]. Furthermore, under the deep learning environment, adversarial perturbation generative network is proposed to preserve image features [12], [13]. However, these privacy protection methods do not protect the FaceID privacy.

What's more, when facial images are released to the third party such as cloud for public downloads, usually some kind of compression technique is used to achieve fast image transmission. The existing methods have to be optimized for the particular image compression framework. One emerging image compression technique is Compressed Sensing (CS), which is an effective technique to increase the image transmission speed through exploring the sparse property of the images [14]. The CS technique can be implemented in both software and hardware, and has already been widely used in optics [15], [16], millimeter wave and terahertz imaging [2], [17], [18], [19], [20], [21], [22], as well as wireless communication [23], [24], [25], [26]. What's more, since the total number of significant sparsifying basis of the images database determines the compression ratio of the CS technique, the facial image basis has to be known to achieve the optimal CS image transmission speed. Also, the facial image database is updated in real time. Thus the sparsifying basis has to be learned adaptively. All of these call for the Dictionary Learning neural Network (DLNet), which can learns the sparsifying basis

This work is partially supported by the National Natural Science Foundation of China under Grant Nos. 61772191, 61976087, 61972058, and 61902123, National Key R&D Projects (2018YFB0704000, 2017YFB0902904), Science and Technology Key Projects of Hunan Province (2019GK2082, 2015TP1004, 2018TP1009, 2018TP2023, 2018TP3001), Transportation Science and Technology Project of Hunan Province (201819), Science and Technology Changsha City (kq1804008, kq2004027). (**Corresponding author: Shaolin Liao.*)

Lu Ou (e-mail: oulu9676@gmail.com), Zheng Qin (e-mail: zqin@hnu.edu.cn) and Dafang Zhang (e-mail: dfzhang@hnu.edu.cn) are with the College of Computer Science and Electronic Engineering, Hunan University, Changsha, 410082, China.

Shaolin Liao (e-mail: liaoshlin@mail.sysu.edu.cn) is with School of Electronics and Information Technology (SEIT) of Sun Yat Sun University (SYSU). He contributes equally as the first author.

Yuan Hong (e-mail: yuan.hong@iit.edu) is with the Department of Computer Science of Illinois Institute of Technology, Chicago, IL 60616, USA.

adaptively in real time [27], [28].

To deal with the FaceID privacy problem under the CS framework, we propose a novel Face-IDentification Privacy (FaceIDP) approach under the CS framework. Without loss of generality, the 2D face identification, instead of the 3D FaceID identification, is used to present the FaceIDP approach. Our major contributions are summarized as follows,

- To our best knowledge, under the CS image transmission application scenario, we propose a novel FaceIDP approach through adding the sanitizing noise at the face encoding coefficients to protect the FaceID privacy, *i.e.*, to prevent the adversary from using users' facial images to breach some FaceID systems.
- To achieve the optimal DP performance, a Dictionary Learning neural Network (DLNet) has been developed to adaptively learn the common dictionary facial image basis of the facial image database so that only sanitizing noise is added to those face encoding coefficients that correspond to the important dictionary facial image basis.
- The Lagrange Multiplier (LM) method has been used to obtain the mathematical formula of the optimized noise scale parameters of the face encoding coefficients for the constrained optimization problem of maximizing the data utility for a given Differential Privacy (DP) budget $\varepsilon$.
- Extensive experiments have been conducted to test the proposed FaceIDP approach with the Labeled Faces in the Wild (LFW) database [29], and the result shows that the FaceIDP approach outperforms other DP approaches without optimization.

## II. Related Work

FaceID as an important technology, has been widely used. Meanwhile, its privacy problem is also very important and challenging. Works have been done on the cryptography-based FaceID problems [1], [10]. These cryptography-based methods can deal with facial image data securely. But the facial image data collection center and the third party need to exchange secrets/keys in a secure channel. It does not fit into our non-interactive setting.

To the best of our knowledge, little research has been conducted on the privacy problem of FaceID protection. However, research on other privacy problems of the facial images have been conducted. For example, to protect image privacy, researchers used pixelization [8] and blurring methods to achieve image obfuscation. Unfortunately, McPherson *et al.*[9] studied pixelization and YouTube face blurring and concluded that the obfuscated images using those methods can be re-identified. Furthermore, in order to deal with such problem, Fan [11] proposed the differentially private pixelization method to protect image features. However, it doesn't focus on differentially private FaceID problem.

Furthermore, regarding the deep learning, Tong and Zheng [12] proposed an adversarial perturbation generative network to generate perturbation to preserve image privacy. Yang *et al.*[13] proposed a facial image privacy protection method by adding perturbation in the principal components of the facial images.

TABLE I
Notations and Definitions

| Symbol | Description |
|---|---|
| $F$ | A facial image. |
| $(A, B)$ | A neighboring facial image pair. |
| $\mathcal{F}$ | $\{F\}$: A FaceID that contains a facial image set of an individual. |
| $(\mathcal{A}, \mathcal{B})$ | Neighboring FaceIDs of a pair of individuals. |
| $F'$ | A noisy facial image. |
| $\overline{F}_M$ | The 1D coding coefficients vector of length $M$. |
| $\mathbb{F}$ | $\{\overline{F}_M\}$: the facial image coding coefficients dataset. |
| $\mathbb{F}_m$ | $\{F_m\}$: the coding coefficient data subset. |
| $\overline{\overline{D}}_{N\times M}$ | The $N \times M$ dictionary basis matrix. |
| $\overline{V}_L$ | The feature vector of a facial image $F$ with length $L$. |
| $\overline{P}_N$ | A 1D pixel vector of a facial image $F$ of length $N$. |
| $\mathbb{P}$ | A set of facial image 1D pixel vectors $\{\overline{P}_N\}$. |
| $f$ | Probability Distribution Function (PDF). |
| $Lap$ | Laplace PDF. |
| $CDF$ | Cumulative distribution function of Laplace distribution. |
| $\Omega$ | Probability space. |
| $Pr$ | Probability over a probability space $\Omega$ |
| $\mathcal{U}$ | Data utility. |
| $\overline{S}$ | Local sensitivity. |
| $\overline{b}_M$ | Laplace noise scale parameter vector of $\overline{F}_M$. |
| $\varepsilon_m$ | Locus privacy budget. |
| $\varepsilon$ | Loci privacy budget. |

Therefore, it is necessary to study the optimal FaceID privacy approach in order to achieve better data utility while still protecting the FaceID system from being attacked by the adversaries, which is the focus of this paper.

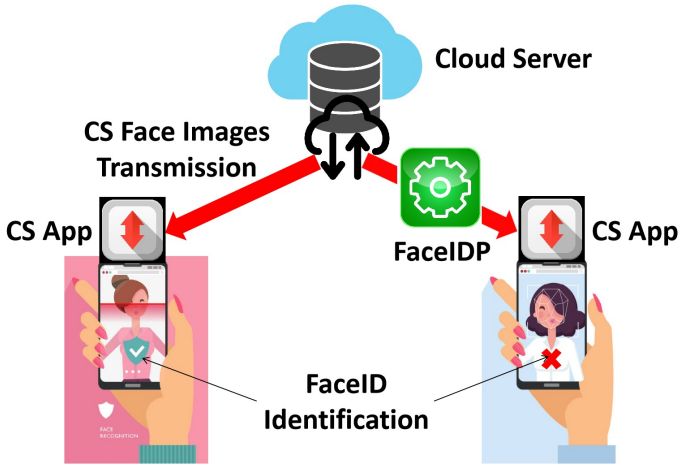


Fig. 1. The FaceIDP problem: the cloud server responds to the smartphone clients' request for face images and starts the CS face images transmission; then the CS app in the smartphone clients receives and reconstructs the face images, which could be used by the adversary to intrude some FaceID identification system (left scenario); so the FaceIDP mechanism runs on the cloud server side to sanitize the face images before the CS face images transmission to protect the FaceID privacy of the face images (right scenario).

## III. Problem Statement

To start, **Table I** lists some key variables used across this paper with their explanations.

### A. System Model

The typical working scenarios of the FaceIDP problem are shown in Fig 1: the smartphone client initiates the request for face images; then the cloud server responds to the request

and sends the CS face images, with or without the FaceIDP protection; upon receiving of the CS face images, the CS app on the smartphone client reconstructs the 3D feature point cloud from the face images, which could be used to intrude some FaceID identification system by the adversary; without the FaceIDP protection, the adversary could break the FaceID identification system; while with the FaceIDP protection, the FaceID identification system is intact.

### B. Facial Images and FaceID

A facial image is any facial image of any individual, which is characterized by its 1D pixel vector denoted as $\overline{P}_N$ of length $N$.

**Definition III.1** (Facial Image Set)**.** *The facial image set $\mathbb{P}$ consists of all individuals' facial images,*

$$\mathbb{P} = \left\{ \overline{P}_N | N = 0, 1, \cdots \right\}. \tag{1}$$

**Definition III.2** (FaceID)**.** *The FaceID $\mathcal{F}$ is the unique identification of an individual: it consists of a set of an individual's facial images $F$,*

$$\mathcal{F} = \left\{ F : \overline{P}_N \right\}. \tag{2}$$

### C. The Facial Images CS

Under the CS framework, sparse facial images can be decomposed in some basis $\overline{\overline{D}}$. So only partial measurements of the image are required for transmission, greatly increasing the image transmission speed.

*1) Facial Image Decomposition:* A facial image can be decomposed into some $N \times M$ dictionary basis $\overline{\overline{D}}_{N\times M}$ with its coding coefficient vector $\overline{F}_M$ of length $M$,

$$\overline{P}_N = \overline{\overline{D}}_{N\times M} \overline{F}_M. \tag{3}$$

*2) The CS Framework:* Assuming the basis $\overline{\overline{D}}$ is known or learned through the DL-NN, the CS performs the following partial measurements,

$$\overline{y}_M = \overline{\overline{\Phi}}_{M\times N} \overline{P}_N = \overline{\overline{\Phi}}_{M\times N} \overline{\overline{D}}_{N\times M} \overline{F}_M, \tag{4}$$

where $\overline{P}_N$ is a facial image pixel vector and an $M \times N$ measurement matrix $\overline{\overline{\Phi}}_{M\times N}$ is used to generate the measurement vector $\overline{y}_M$ of length $M$. Also, the $N \times M$ dictionary basis matrix $\overline{\overline{D}}_{N\times M}$ has a coefficient coding vector $\overline{F}_M$ of length $M$.

### D. Facial Image Data

*1) Facial Image Data Records:* A facial image data record is a set of facial images of an individual defined as **Definition** III.3.

**Definition III.3** (Data Record)**.** *A data record of the FaceIDP is for all facial images $F$ of an individual, identified by the FaceID $\mathcal{F}$ and represented by the coding coefficients vector $\overline{F}_M$,*

$$\left\{ \overline{F}_M \right\} : F \in \mathcal{F}. \tag{5}$$

*2) FaceID Datasets:* A FaceID dataset of the FaceIDP consists of facial image sets of all users,

**Definition III.4** (Dataset)**.** *The dataset of the FaceIDP is given by,*

$$\mathbb{F} = \left\{ \mathcal{F} \right\}. \tag{6}$$

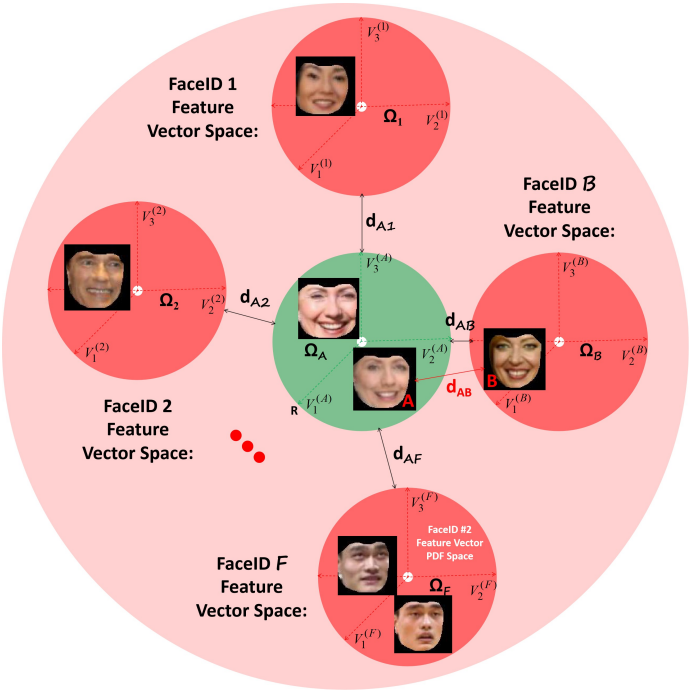


Fig. 2. Feature vector spaces for FaceID $\mathcal{A}$: all FaceIDs other than FaceID $\mathcal{A}$ are neighboring face images sets of FaceID $\mathcal{A}$; also, the closest neighboring FaceID is the FaceID with the minimum separation distance $d$ from the FaceID $\mathcal{A}$, *i.e.*, FaceID $\mathcal{B}$; Finally, the closest face images pair is the face images in the closest FaceID pair that has the minimum separation distance, *i.e.*, the face pair $(A, B)$.

*3) The Neighboring Data Records:* The neighboring data records are a user's facial image set and all other users' facial image set, represented by their coding coefficients vectors,

**Definition III.5** (The Neighboring Facial Image Sets)**.** *The neighboring facial image sets are the face image set of a user ($A \in \mathcal{A}$) and the face image set of all other users ($B \in \mathcal{B}$),*

$$\left( \left\{ \overline{F}_M^{(A)} : A \in \mathcal{A} \right\}, \left\{ \overline{F}_M^{(B)} : B \notin \mathcal{B} \right\} \right) : \mathcal{B} \cup \mathcal{A} = \varnothing. \tag{7}$$

*4) The Closest Neighboring Face Image Pair:* The closest neighboring face images are a pair of most similar face images of the neighboring face image sets in **Definition III.5**, under the measure of $\mathcal{M}$,

**Definition III.6** (The Closest Neighboring Face Images)**.** *The cloest neighboring face image pair is,*

$$(A, B) : \underset{B}{argmin} \left\{ |\mathcal{M}(A) - \mathcal{M}(B)| \right\}. \tag{8}$$

### E. The FaceIDP under the CS Framework

In this paper, we study the privacy problem of the FaceID under the CS framework: to protect the user's FaceID from being attacked by the adversary while still providing the

optimal data utility for the face image CS application to achieve fast face image transmission.

### F. The DP Framework

**Definition III.7** (Differential Privacy)**.** *Let $\mathcal{M}$ be a randomized query measure function, $\mathcal{A}$ be any output of $\mathcal{M}$, $A$ and $B$ be two neighboring datasets. $\mathcal{M}$ will be $\varepsilon$-differential private, if the following is satisfied*

$$exp(-\varepsilon) \leq \frac{Pr(\mathcal{M}(A) \in \mathcal{A})}{Pr(\mathcal{M}(B) \in \mathcal{A})} \leq exp(\varepsilon).$$

Under the scenario of FaceID, we are interested only in whether the FaceID system can distinguish between two closest face images, the neighboring datasets are the the closest neighboring face images given in **Definition III.6**.

*1) The FaceID Query Measure Function:* The FaceID query measure function $\mathcal{M}$ is to identify a user's identification based on all noisy face images of the user. The measured outcomes are either "1" if the measure function $\mathcal{M}$ can identify the user's identification or "0" if the measure function $\mathcal{M}$ cannot identify the user's identification, above some confidence probability level.

### G. Adversary Model

The goal of the FaceIDP is to prevent the adversary from using a set of sanitized facial images of a user to pass the FaceID identification system. Thus the sanitizing noise has to be added in such a way that the FaceID identification system cannot distinguish if a set of noisy facial images belong to a user or not, with some confidence probability level.

## IV. Concepts

In this section, some basic concepts are given.

### A. The Locus Differential Privacy

The Locus Differential Privacy (Locus-DP) is defined at each locus of a loci [4], [5], [6],

**Definition IV.1** (The Locus Differential Privacy)**.** *The random obfuscating measure function $\mathcal{M}'$ for a data subset $F_m$ of a given data set $\overline{F}_M$: $F_m \in \overline{F}_M, m = 1, \cdots M$, is said to be $\varepsilon_m$-deferentially private if the following probability condition is satisfied after the sanitizing noise is added to $F_m$,*

$$\exp(-\varepsilon_m) \leq \frac{Pr\left\{\mathcal{M}'\left(F_m^{(i)}\right)\right\} = F'_m}{Pr\left\{\mathcal{M}'\left(F_m^{(j)}\right)\right\} = F'_m} \leq \exp(\varepsilon_m),$$

*where in this paper, the random obfuscating measure function $\mathcal{M}'$ is the Euclidean norm distance measure $\mathcal{M}$ given below that is sanitized by the Laplace noise,*

$$\mathcal{M}\left(\overline{F}_M^{(A)} - \overline{F}_M^{(B)}\right) \equiv \left|\left|\overline{F}_M^{(A)} - \overline{F}_M^{(B)}\right|\right|_2 = \sqrt{\sum_{\ell=1}^{L}\left(F_m^{(A)} - F_m^{(B)}\right)^2}. \quad (9)$$

### B. The Loci DP or DP

The loci DP or simply DP [30], [31] is defined on the data set $\overline{F}_M : \{F_m | m = 1, \cdots, M\}$ in such as way that the sanitizing noise added to $F_m$ according to **Definition III.7** will satisfy the following probability relation,

**Definition IV.2** (Loci DP or DP)**.**

$$\exp(-\varepsilon) \leq \frac{Pr\left\{\mathcal{M}\left(\overline{F}_M\right)\right\} = \overline{F}'_M}{Pr\left\{\mathcal{M}\left(\overline{F}_M\right)\right\} = \overline{F}'_M} \leq \exp(\varepsilon).$$

### C. Joint Probability Bounds

From **Definition IV.2**, it is clear that the privacy budget $\epsilon$ is closely related to the lower and upper bounds of numerator and denominator. The joint Probability Distribution Function (PDF) of multivariate random variables vector $\overline{F}_M$ of length $M$, denoted as $f(\overline{F}_M)$, has its lower and upper bounds on a domain $\Omega$ given as follows,

**Lemma IV.1** (Bounds of the Joint Probability)**.** *The lower and upper bounds of the joint probability of $f(\overline{F}_M)$ on a domain $\Omega$ is given by*

$$Pr(\overline{F}_N \in \Omega) \begin{cases} \geq \max\limits_{\Omega_{F_m}} \left\{ \prod\limits_{m=1}^{M} Pr(X_m \in \Omega_{F_m}) \right\}, \\ \leq Pr(\overline{F}_M \in \Omega) \leq \min_m \{Pr(F_m \in \Omega)\}, \end{cases}$$

*where $\overline{\Omega} = \overline{I} - \Omega$ is the complementary domain with $\overline{I}$ being the entire domain of interest; and $\Omega_{F_m}$ is the sub-domain in which all $\overline{F}_M$ belongs to $\Omega$.*

***Proof:*** The joint probability distribution can be expressed in terms of the conditional probability distribution,

$$f(\overline{F}) = f(F_m) f(\cdots F_{m-1}, F_{m+1}, \cdots | F_m) \leq f(F_m),$$

where the following conditional probability property has been used,

$$f(\cdots F_{m-1}, F_{m+1}, \cdots | F_m) \leq 1.$$

from which the probability in domain $\Omega$ is given by,

$$Pr(\overline{F} \in \Omega) = \int_{F_1} \cdots \int_{F_M} f(\overline{F}) dF_1 \cdots dF_M \quad (10)$$

$$\leq \int_{F_1} \cdots \int_{F_M} f(F_m) dF_1 \cdots dF_M = Pr(F_m \in \Omega),$$

and the upper bound on the right hand side of **Lemma IV.1** is proved,

$$Pr(\overline{F} \in \Omega) \leq \min_m \{Pr(F_m \in \Omega)\}. \quad (11)$$

The lower bound of the left hand side of **Lemma IV.1** can be obtained by finding the sub-domains of all $F_n$, denoted as $\Omega_{F_m}$, in which all $\overline{F}_M$ belongs to $\Omega$ and the probability is given by,

$$Pr(\overline{F}_M \in \overline{\Omega}) \leq \min_m \left\{Pr(F_m \in \overline{\Omega})\right\}, \quad (12)$$

from which the lower bound of the probability in domain $\Omega$ is given by,

$$Pr(\overline{F} \in \Omega) \geq \int_{F_1} \cdots \int_{F_N} f(\overline{F}) dF_1 \cdots dF_M$$
$$\geq \int_{\Omega_{F_1}} \cdots \int_{\Omega_{F_M}} f(F_m) dF_1 \cdots dF_M = \prod_{n=1}^{M} Pr(F_m \in \Omega_{F_m}),$$

where independence has been assumed for all elements of $\overline{F}_M$ and the lower bound is thus obtained as,

$$Pr(\overline{F} \in \Omega) \geq \max_{\Omega_{F_m}} \left\{ \prod_{m=1}^{M} Pr(F_m \in \Omega_{F_m}) \right\},$$

from which **Lemma IV.1** is proved. □

### D. Data Utility in the CS Framework

When the coding coefficients noise $\overline{n}_M$ is added to a face image's coding coefficients $\overline{F}_M$, the noisy image is thus obtained as,

$$\overline{P}'_N = \overline{P}_N + \overline{\overline{D}}_{N\times M} \overline{n}_M. \tag{13}$$

Then, the data utility under the CS framework is thus defined as follows,

**Definition IV.3** (Data Utility)**.** *The data utility is defined as the visual quality of the image [12]: here the expectation of the variance of the reconstructed noisy image from the original image,*

$$\mathcal{U} = E\left\{ \left| \overline{P}'_N - \overline{P}_N \right|_2^2 \right\}. \tag{14}$$

Substituting Eq. (13) into Eq. (14), the data utility is obtained,

$$\mathcal{U} = \mathrm{E}\left\{ \overline{n}_M^T \left[ \overline{\overline{D}}_{N\times M}^T \overline{\overline{D}}_{N\times M} \right] \overline{n}_M \right\}$$
$$= \sum_{m=1}^{M} W_m \sigma_m^2, \quad W_m = \sum_{n=1}^{N} D_{n,m}^2, \tag{15}$$

where $\sigma_m$ is the standard deviation of the noise component $n_m$, which is assumed to be independent from each other.

### E. Differential Privacy for FaceID

FaceID can be considered as a special classification method. FaceID takes the human face as input and extract the feature vector, denoted by $\overline{F}$, of the human face and then adopts some judgement metrics, denoted by $\mathcal{M}$, to identify which individual the human face will be assigned to, denoted by $\mathcal{A}$. The privacy budget $\varepsilon$ for FaceID is defined as follows,

**Definition IV.4** (Privacy Budget)**.** *The privacy budget for the DP problem of FaceID is defined as the negative logarithmic value of the maximum ratio between the probabilities of the assigned FaceID $\mathcal{A}$ of any pair of human face images $(A, B)$ that belong to two different FaceIDs $A \in \mathcal{A}, B \in \mathcal{B}$ after the sanitizing noise is added,*

$$\varepsilon = -\ln \left( \max \left\{ \frac{Pr\left(\overline{P}_N^{(B')} : B' \in \mathcal{A}\right)}{Pr\left(\overline{P}_N^{(A')} : A' \in \mathcal{A}\right)} \right\} \right), \tag{16}$$

*where $\overline{P}_N^{(B')}$ is the noisy face image $B'$ that belongs to the FaceID $\mathcal{B}$ and $\overline{P}_N^{(A')}$ is the noisy face image $A'$ that belongs to the FaceID $\mathcal{A}$.*

### F. Optimal Noise Scale Parameters

The data utility in **Definition IV.3** and the privacy budget in **Definition 16** are a balanced pair: if the data utility is high ($\mathcal{U}$ is low), the privacy is low ($\varepsilon$ is high) and vice versa. Also they are both functions of the noise scale parameter of $\overline{n}_M$, denoted as $\overline{b}_M$. So it is desire to optimize the data utility $\mathcal{U}$ for the given privacy budget $\varepsilon = \varepsilon_0$, which is the constraint optimization problem,

**Lemma IV.2.** *The constraint optimization of the data utility $\mathcal{U}$ for a given privacy budget $\varepsilon$ can be done through the Lagrange Multiplier (LM) method,*

$$\frac{\partial}{\partial \overline{b}_M} \mathcal{L}(\overline{b}_M) = 0; \quad \varepsilon(\overline{b}_M) = \varepsilon_0, \tag{17}$$
$$\mathcal{L}(\overline{b}_M) = \mathcal{U}(\overline{b}_M) + \lambda \left[ \varepsilon(\overline{b}_M) - \varepsilon_0 \right],$$

***Proof:*** The constraint optimization problem can be expressed as follows,

$$\min \left\{ \mathcal{U}(\overline{b}) \right\}, \quad s.t. \quad \varepsilon(\overline{b}) = \varepsilon_0, \tag{18}$$

The optimization is achieved when the derivatives of $\mathcal{L}(\overline{b}_N)$ with respect to both $\overline{b}_N$ and the Lagrange multiplier $\lambda$ and **Lemma IV.2** is proved. □

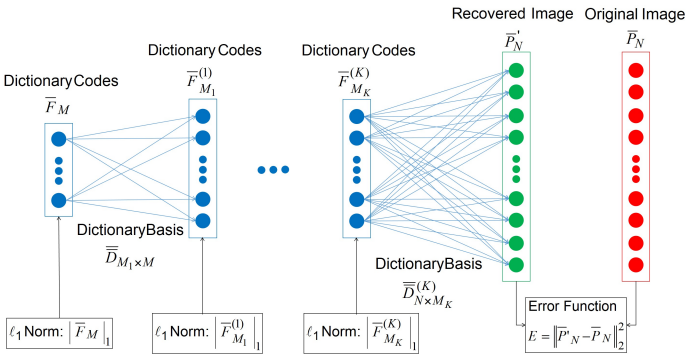


Fig. 3. The working principle of the DLNet to learn the sparse dictionary basis of the face images.

## V. Our Approach

In this section, we present the details of our proposed FaceIDP approach. Under the CS framework, the face images are represented as the coding coefficients vectors under some face images' dictionary basis. So we first build the DLnet to learn the face images' dictionary basis. Then the sanitizing

noise is added to the coding coefficients vectors. After that, the probabilities of the noisy neighboring face image pair $(A, B)$ that are assigned the FaceID $\mathcal{A}$ are calculated. Finally, the privacy budget's lower and upper bounds are obtained.

### A. The DLNet

Fig. 3 shows the working principle of the DLNet for face images' dictionary basis learning [28]: the DLNet consists of multiple layers with their basis denoted as $\overline{\overline{D}}^{(k)}_{M_{i+1}\times M_k}, k = 1, K$ and the corresponding coding coefficients denoted as $\overline{\overline{F}}^{(k)}_{M_k}$; Both the dictionary basis and the coding coefficients can be trained through minimizing the two error functions, *i.e.*, the mean square error of the reconstructed image $E$,

$$E = \left|\overline{P}'_N - \overline{P}_N\right|^2, \tag{19}$$

as well as the $\ell_1$ norm of the sparse codes $\left|\overline{\overline{F}}^{(k)}_{M_k}\right|_1$.

The DLNet is trained through two sequential steps: 1) updating of the parameters through the Stochastic Gradient Descent (SGD) method; and 2) performing the $\ell_1$ norm operation on the updated parameters.

*1) The SGD Updating:* First, the gradient of parameter $x$, denoted as $\nabla_x E$, can be obtained through the train rule,

$$\nabla_x E = -\sum_{n=1}^{N}\left(P_n - P'_n\right)\nabla_x P'_n, \tag{20}$$

and the parameter $x$ is updated as follows

$$x = x - \eta\nabla_x E, \tag{21}$$

with $\eta$ being the learning rate and the parameter $x$ is either the dictionary bases or the coding coefficients,

$$x = \left\{\overline{\overline{D}}_{M_k\times M}, \overline{\overline{F}}^{(k)}_{M_k}\right\}. \tag{22}$$

*2) The $\ell_1$-norm Operation:* Then, the $\ell_1$-norm Operation is performed on the SGD updated coding coefficients $\overline{\overline{F}}^{(k)}_{M_k}$ through the Iterative Soft Thresholding Algorithm (ISTA) to achieve the sparsity of the coding coefficients,

$$\overline{\overline{F}}^{(k)}_{M_k} = sign\left\{\overline{\overline{F}}^{(k)}_{M_k}\right\}\max\left\{0, \overline{\overline{F}}^{(k)}_{M_k} - \lambda\right\}, \tag{23}$$

where $\lambda$ is the thresholding value.

Finally, after the training of the DLNet, the total dictionary basis $\overline{\overline{D}}_{N\times M}$ is obtained as follows,

$$\overline{\overline{D}}_{N\times M} = \overline{\overline{D}}_{N\times M_K}\left(\prod_{k=1}^{K}\overline{\overline{D}}_{M_{k+1}\times M_k}\right)\overline{\overline{D}}_{M_1\times M}. \tag{24}$$

### B. Data Utility

For joint Laplace distribution of $\overline{F}_M$, the data utility $\mathcal{U}$ in **Definition IV.3** is reduced to the following,

$$\mathcal{U} = \sum_{m=1}^{M} 2W_m b_m^2, \tag{25}$$

where the variance of the Laplace distribution $\sigma_m^2 = 2b_m^2$ has been used.

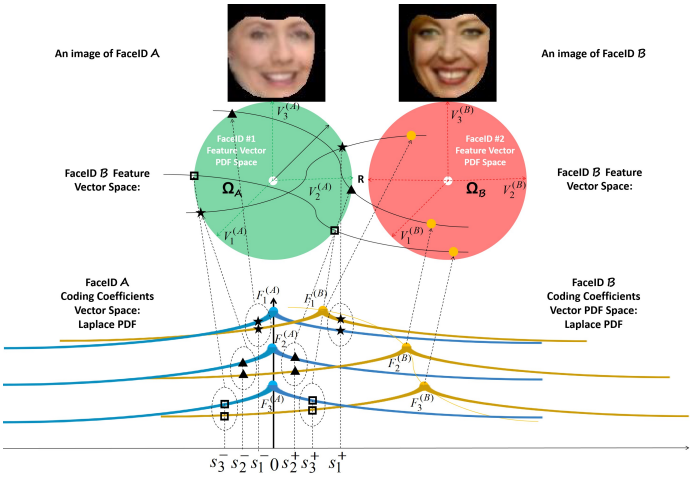


Fig. 4. PDF spaces and sensitivities of the FaceIDP shows the relation between the PDF space of the feature vector and the PDF space of the coding coefficients vector for the closest neighboring face image pair ($A \in \mathcal{A}, B \in \mathcal{B}$). Both the origin of the feature vector space and the coding coefficients space are set to the center of FaceID $\mathcal{A}$.

### C. Privacy Budget

According to the definition of the privacy budget in **Definition 16**, the lower bound and upper bound of two probabilities have to be computed: 1) the probability that a noisy face image $B'$ of FaceID $\mathcal{B}$ is mistakenly assigned to FaceID $\mathcal{A}$; and 2) the probability that a noisy image $A'$ of FaceID $A \in \mathcal{A}$ is still assigned the correct FaceID $\mathcal{A}$. These probabilities are related to two inter-correlated provability spaces, as shown in Fig. 4: 1) the FaceID feature vector space; and 2) the face coding coefficients space. First, the FaceID $\mathcal{A}$ is assigned to a face image $F$ through the Euclidean norm measure $\mathcal{M}$ on the feature vector $\overline{V}_L$ of length $L$,

$$F \in \mathcal{A} : \mathcal{M}\left(\overline{V}_L\right) = \left|\left|\overline{V}_L\right|\right|_2 \in \Omega_{\mathcal{A}} = \left|\left|\overline{V}_L\right|\right|_2 \le R,$$
$$\left|\left|\overline{V}_L\right|\right|_2 \equiv \sqrt{\sum_{\ell=1}^{L}\left(V_L\right)^2}, \tag{26}$$

where $R$ is the radius of FaceID $\mathcal{A}$ as shown in Fig. 4.

Then, the probability of a face $F$ assigned FaceID $\mathcal{A}$ is given by,

$$P_F = Pr\left(\mathcal{M}\left\{\overline{V}_L\right\} \in \Omega_{\mathcal{A}}\right). \tag{27}$$

Also, the feature vector space $\overline{V}_L$ is related to the face coding coefficients vector $\overline{F}_M$. For example, when the change of a single face coding coefficients element $F_m$ corresponds to a provability curve $\left|\left|\overline{V}_L\right|\right|_2$ in the feature vector space $\overline{V}_L$, as shown in Fig. 4.

**Definition V.1** (Probability Boundary Edges)**.** *The probability boundary edges define the probability space within which the noisy image $B'$ is assigned the FaceID $\mathcal{A}$, while other coefficients elements are set to zeros, i.e., $F^{(B')}_{m'} = 0, m' \neq m$,*

$$\left(s_m^-, s_m^+\right) \equiv F_m^{(B')} \in \left(s_m^-, s_m^+\right) \in \Omega_{\mathcal{A}}. \tag{28}$$

*1) Probability of the Noisy Face $B$ Assigned FaceID $\mathcal{A}$:* The probability that a noisy face image from its original face image $B$ of FaceID $\mathcal{B}$ is mistakenly assigned to FaceID $\mathcal{A}$ is given by,

$$\begin{aligned} P_B &\equiv Pr\left(\overline{F}_M^{(B')} : B' \in \mathcal{A} \middle| \bar{b}_M\right) = Pr\left(\mathcal{M}\left\{\overline{F}_M^{(B')}\right\} \in \Omega_{\mathcal{A}}\right) \\ &= \int\limits_{\mathcal{M}\left\{\overline{F}_M^{(B')}\right\} \in \Omega_{\mathcal{A}}} \cdots \int f_{\bar{b}_M}\left(F_1^{(B')}, \cdots, F_M^{(B')}\right) d\overline{F}_M^{(B')}, \end{aligned}$$

where $f_{\bar{b}_M}$ is the joint Laplace PDF of $\overline{F}_M^{(B')}$ with the noise scale parameter vector of $\bar{b}_M$,

$$\begin{aligned} f_{\bar{b}_M}\left(F_1^{(B')}, \cdots, F_M^{(B')}\right) &= Lap\left(\overline{F}_M^{(B')} \middle| \bar{b}_M\right), \qquad (29) \\ Lap\left(\overline{F}_M^{(B')} \middle| \bar{b}_M\right) &= \prod_{m=1}^{M} Lap\left(F_m^{(B')} \middle| b_m\right), \end{aligned}$$

where independence has been assumed for $\overline{F}_M$.

Now look at the lower bound and upper bound of the probability according to **Lemma IV.1**. First, the probability upper bound is given by,

$$\begin{aligned} P_B^+ &\equiv \max\left\{Pr\left(\mathcal{M}\left\{\overline{F}_M^{(B')}\right\} \in \Omega_{\mathcal{A}}\right)\right\} \qquad (30) \\ &= \min_m \int\limits_{F_m \in \Omega_{\mathcal{A}}} \cdots \int Lap\left(F_m^{(B')} - F_m^{(B)} \middle| \bar{b}_m\right) dF_m^{(B')}, \\ &= \min_m \left\{CDF\left(s_m^+ - F_m^{(B)}\right) - CDF\left(s_m^- - F_m^{(B)}\right)\right\}, \end{aligned}$$

where $CDF$ is the cumulative distribution function of the Laplace distribution; and $s_m^-$ and $s_m^+$ are the left and right probability boundary edges of coding coefficients element $m$ in $\Omega_{\mathcal{A}}$ given in **Definition V.1**.

Similary, according to according to **Lemma IV.1**, the probability lower bound is given by,

$$\begin{aligned} P_B^- &\equiv \min\left\{Pr\left(\mathcal{M}\left\{\overline{F}_M^{(B')}\right\} \in \Omega_{\mathcal{A}}\right)\right\} \\ &= \max_{\Omega_{\mathcal{A},m}} \left\{\prod_{m=1}^{M} Pr\left(F_m^{(B')} \in \Omega_{\mathcal{A},m}\right)\right\}, \qquad (31) \end{aligned}$$

where the local probability space $\Omega_{\mathcal{A},m}$ can be obtained as follows,

**Definition V.2** (Local Probability Domain)**.** *The local probability domain is defined as the maximum linear scaling of space bounded by the probability boundary edges such that the noisy image $B'$ is assigned the FaceID $\mathcal{A}$,*

$$\Omega_{\mathcal{A},m} = \alpha\left(s_m^-, s_m^+\right) : \alpha = \underset{\alpha}{argmax}\left\{B' \to \mathcal{A}\right\}, \qquad (32)$$

*for all coding coefficients elements $m = 1, \cdots M$ and $\alpha$ is the linear scaling parameter.*

Now the probability lower bound in Eq. (31) reduces to the following,

$$P_B^- = \prod_{m=1}^{M} \left\{CDF\left(\alpha s_m^+ - F_m^{(B)}\right) - CDF\left(\alpha s_m^- - F_m^{(B)}\right)\right\}. \qquad (33)$$

*2) Probability of the Noisy Face $A$ Assigned FaceID $\mathcal{A}$:* Similarly, the probability that a noisy image $A'$ from a FaceID $A \in \mathcal{A}$ is still assigned the correct FaceID $\mathcal{A}$ has the following uppper bound and lower bound,

$$\begin{aligned} P_A &\equiv Pr\left(\mathcal{M}\left\{\overline{F}_M^{(A')}\right\} \in \Omega_{\mathcal{A}}\right) \\ P_A^+ &\equiv \max\left\{Pr\left(\mathcal{M}\left\{\overline{F}_M^{(A')}\right\} \in \Omega_{\mathcal{A}}\right)\right\} \\ &= \min_m \left\{CDF\left(s_m^+\right) - CDF\left(s_m^-\right)\right\}, \\ P_A^- &\equiv \min\left\{Pr\left(\mathcal{M}\left\{\overline{F}_M^{(A')}\right\} \in \Omega_{\mathcal{A}}\right)\right\} \\ &= \prod_{m=1}^{M} \left\{CDF\left(\alpha s_m^+\right) - CDF\left(\alpha s_m^-\right)\right\}. \qquad (34) \end{aligned}$$

*3) Privacy Budget Bounds:* With the above probability bounds, the privacy budget bounds can be obtained.

**Lemma V.1.** *The privacy budget has the lower bound and upper bound of*

$$\varepsilon^-\left(\bar{b}_M\right) \le \varepsilon\left(\bar{b}_M\right) \le \varepsilon^+\left(\bar{b}_M\right), \qquad (35)$$

*where $\varepsilon^-$ and $\varepsilon^+$ are the lower bound and upper bound given below.*

***Proof:*** The privacy budget $\varepsilon$ is obtained from **Definition 16**,

$$\varepsilon\left(\bar{b}_M\right) = -\ln\left(\max_{(A,B)}\left\{\frac{P_B}{P_A}\right\}\right). \qquad (36)$$

From Eq. (30) and Eq. (34), the lower bound of the privacy budget is given by,

$$\begin{aligned} \varepsilon^- &= \max_{(A,B)}\left\{\ln\left(\frac{P_A^-}{P_B^+}\right)\right\} \\ &= \max_{(A,B)}\left\{\ln\left(\frac{P_A^-}{\min\left\{P_B^{+o}, P_B^{+i}\right\}}\right)\right\}, \qquad (37) \end{aligned}$$

where

$$\begin{aligned} P_A^- &= \prod_m \left\{1 - \frac{\exp\left(-\frac{\alpha s_m^+}{b_m}\right) + \exp\left(\frac{\alpha s_m^-}{b_m}\right)}{2}\right\}, \\ P_B^{+o} &= \min_{F_m^B \notin \left(s_m^-, s_m^+\right)} \left\{\frac{\left|\exp\left(-\frac{S_m^+}{b_m}\right) - \exp\left(-\frac{S_m^-}{b_m}\right)\right|}{2}\right\}, \\ P_B^{+i} &= 1 - \max_{F_m^B \in \left(s_m^-, s_m^+\right)} \left\{\frac{\exp\left(-\frac{S_m^+}{b_m}\right) + \exp\left(-\frac{S_m^-}{b_m}\right)}{2}\right\}, \end{aligned}$$

and $S_m^-$ and $S_m^+$ are the distances from the left and right probability boundary edges given below,

$$S_m^- = \left|F_m^B - s_m^-\right|; \quad S_m^+ = \left|F_m^B - s_m^+\right|. \qquad (38)$$

Similarly, from Eq. (33) and Eq. (34), the upper bound of the privacy budget is given by,

$$\varepsilon^+ = \max_{(A,B)}\left\{\ln\left(\frac{P_A^+}{P_B^{-i} P_B^{-o}}\right)\right\}, \qquad (39)$$

where

$$P_A^+ = \min_m \left\{ 1 - \frac{\exp\left(-\frac{S_m^+}{b_m}\right) + \exp\left(-\frac{S_m^-}{b_m}\right)}{2} \right\}, \quad (40)$$

$$P_B^{-o} = \prod_{F_m^B \notin \left(s_m^-, s_m^+\right)} \left\{ \frac{\left| \exp\left(-\frac{\tilde{S}_m^+}{b_m}\right) - \exp\left(-\frac{\tilde{S}_m^-}{b_m}\right) \right|}{2} \right\},$$

$$P_B^{-i} = \prod_{F_m^B \in \left(s_m^-, s_m^+\right)} \left\{ 1 - \frac{\exp\left(-\frac{\tilde{S}_m^+}{b_m}\right) + \exp\left(-\frac{\tilde{S}_m^-}{b_m}\right)}{2} \right\},$$

where $\tilde{S}_m^+$ and $\tilde{S}_m^-$ are defined as follows,

$$\tilde{S}_m^+ = \left| F_m^B - \alpha s_m^+ \right|, \quad \tilde{S}_m^- = \left| F_m^B - \alpha s_m^- \right|.$$

After some mathematics calculation, the upper bound of the privacy budget can be expressed as follows,

$$\varepsilon^+ = \delta + \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m},$$

with

$$\delta = \max_{(A,B)} \left\{ \ln \left( \frac{P_A^+}{\prod P_B^{-i} \prod \tilde{P}_B^{-o}} \right) \right\}$$

$$\tilde{P}_B^{-o} = \prod_{F_m^B \notin \left(s_m^-, s_m^+\right)} \left\{ \frac{\left| 1 - \exp\left(-\frac{\left|\tilde{S}_m^+ - \tilde{S}_m^-\right|}{b_m}\right) \right|}{2} \right\},$$

where the local sensitivity $S_m$ is defined as follows,

**Definition V.3** (Local Sensitivity)**.** *The local sensitivity is defined as closest distance from the coding coefficients elements to their probability boundary edges,*

$$S_m = \min \left\{ \tilde{S}_m^-, \tilde{S}_m^+ \right\}. \quad (41)$$

□

## VI. Analysis

In this section, the FaceIDP noise mechanism is proved to satisfy the $\varepsilon$-deferentially private guarantee.

**Theorem VI.1.** *The noise mechanism of the FaceIDP satisfies $\varepsilon$-differential privacy,*

$$\exp(-\varepsilon) \leq \frac{Pr\left(\overline{F}_M^{(B')} : B \in \mathcal{A}\right)}{Pr\left(\overline{F}_M^{(A')} : A \in \mathcal{A}\right)} \leq \exp(\varepsilon),$$

*with*

$$\varepsilon = \delta + \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m}.$$

***Proof:*** From **Lemma V.1**, the privacy budget has its upper bound of

$$\varepsilon\left(\bar{b}_M\right) \leq \delta + \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m},$$

from which **Theorem VI.1** is proved. □

## VII. Optimal Noise for Better Data Utility

From **Lemma IV.2**, the data utility $\mathcal{U}$ can be optimized to obtain the optimal noise scale parameter $\bar{b}_M$, for a given privacy budget $\varepsilon$,

$$\min_{\bar{b}_M} \left\{ \mathcal{U} = \sum_{m=1}^{M} 2W_m b_m^2 \,\middle|\, \varepsilon(\bar{b}_M) = \varepsilon_0 \right\}. \quad (42)$$

With the probabilities given in Eq. (40), the LM optimization problem in Eq. (42) can be solved numerically. Under the approximation that $P_A^+$, $P_B^{-i}$ and $P_B^{-i}$ are constants, the privacy budget factor $\delta$ is also a constant and an effective privacy given budget $\varepsilon_0'$ can be defined according to **Theorem VI.1**

$$\varepsilon_0' = \varepsilon_0 - \delta = \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m}, \quad (43)$$

and the LM optimization problem in Eq. (42) reduces to the following,

$$\frac{\partial}{\partial \bar{b}_M} \mathcal{L}(\bar{b}_M) = 0; \quad \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m} = \varepsilon_0' = \varepsilon_0 - \delta, \quad (44)$$

$$\mathcal{L}(\bar{b}_M) = \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} 2W_m b_m^2 + \lambda \left[ \sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m}{b_m} - \varepsilon_0' \right].$$

**Theorem VII.1** (Optimal Noise Scale Parameters)**.** *The optimal noise scale parameters vector $b_m^*$ is given by,*

$$b_m^* = \frac{S_m}{\varepsilon_m'}, \quad (45)$$

*with*

$$\varepsilon_m' = p_m \varepsilon_0'$$

$$p_m = \frac{W_m^{1/3} S_m^{2/3}}{\sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} W_m^{1/3} S_m^{2/3}}.$$

***Proof:*** From Eq. (44),

$$\frac{\partial}{\partial \bar{b}_M} \mathcal{L}(\bar{b}_M) = 0 \rightarrow b_m = \left( \frac{\lambda S_m}{4W_m} \right)^{1/3}, \quad (46)$$

from which the constraint of the privacy budget is given by,

$$\sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} \frac{S_m^{2/3} (4W_m)^{1/3}}{(\lambda)^{1/3}} = \varepsilon_0', \quad (47)$$

and

$$\lambda = \left( \frac{\sum_{F_m^B \notin \left(s_m^-, s_m^+\right)} S_m^{2/3} (4W_m)^{1/3}}{\varepsilon_0'} \right)^3. \quad (48)$$

Substituting Eq. (48) into Eq. (46), the noise scale parameters are obtained and **Theorem VII.1** is proved. □

## VIII. Algorithm

The algorithm for the FaceIDP is shown in Algorithm 1.

**Algorithm 1** FaceIDP

**Input:** Face images $\mathbb{P}$ and privacy budget $\varepsilon$.
**Output:** Sanitized face images $\mathbb{P}'$ satisfying DP.

$\mathbb{P}' = \varnothing$
Learn the sparse dictionary basis $\mathbb{D}$ of the face images data set through the DLNet.
**for** each face image $\overline{P}_N \in \mathbb{P}$ **do**
    Decompose the face image $\overline{P}_N$ into the product of the selected dictionary basis $\overline{\overline{D}}_{N\times M}$ and coding coefficients vector $\overline{F}_M$.
    Compute the weight vector $\overline{W}_M$ according to Eq. (15).
    Calculate the sensitivity vector $\overline{S}_M$ according to Eq. (41).
    Compute the optimal noise scale parameters $\overline{b}_M$ according to **Theorem VII.1**.
    Obtain the coding coefficients noise through the joint Laplace distribution: $\overline{\delta}_M = \prod_{m=1}^{M} Lap(F_m|b_m)$.
    Obtain the sanitized noisy image $\overline{P}'_N$ according to Eq. (13).
    Update the sanitized image dataset: $\mathbb{P}' = \mathbb{P}' \cup \overline{P}'_N$.
**return** $\mathbb{P}'$.

## IX. Experimental Results

During the FaceIDP experiment, the pre-trained model of Dlib, a ResNet based neural network, is used in Python 3.7 to perform the face recognition. The has been trained and tested with two databases: 1) LFW database [29] and 2) PubFig database [32]. On one hand, LFW is a database of face photographs designed for studying the problem of unconstrained face recognition. On the other hand, unlike most other existing face databases, these images of the PubFig database are taken in completely uncontrolled situations with non-cooperative subjects.

The face recognition consists of 4 common stages: face detection, face align, face encodings representation and face verification. The facial landmark detector file is "shape_predictor_68_face_landmarks.dat" and the ResNet model file is "dlib_face_recognition_resnet_model_v1.dat" [33]. The face encoding feature vector $\overline{V}_L$ has a dimension of $L = 128$ and the Euclidean distance is used to recognize the faces with a threshold of 0.6.

To show the efficiency of our optimal FaceIDP method, we compared it to the standard-DP method and the partial-DP method where sanitizing noise is added to partial coding coefficients that lie outside of the Probability Boundary Edges according to **Definition V.1** : $F_m^B \notin (s_m^-, s_m^+)$, *i.e.*, sanitizing noise is added to coding coefficients that have the most significant effect on the face encoding feature vectors, as shown in **Theorem VI.1**.

### A. The DLNet

First, the common bases of the face images $\overline{\overline{D}}_{N\times M}$ are learned through the DLNet in **Section V-A**. 1000 face images of the LFW database are used to train the DLNet to obtain 100 face dictionary bases. During the training, the learning rate of the SGD $\eta$ and the ISTA thresholding value $\lambda$ are set as follows,

$$\eta = 0.01; \quad \lambda = 0.01 \max\left\{\overline{\overline{F}}^{(k)}_{M_k}\right\}, \quad k = 1, 2, \cdots, K.$$

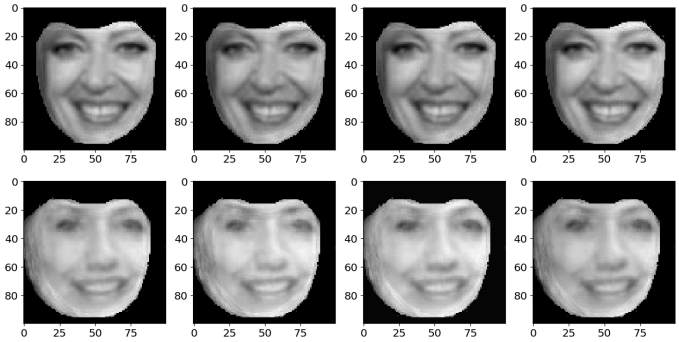

Fig. 5. (LFW Database) The closest neighboring Face $A$ and Face $B$ for a given data utility $\mathcal{U} = 13$: 1st column) original faces; 2nd column) Standard-DP noisy faces; 3rd column) Partial-DP noisy faces; and 4th column) the optimal FaceIDP noisy faces.

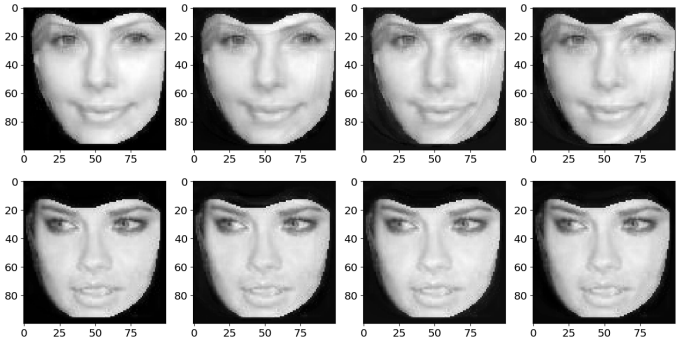

Fig. 6. (PubFig Database) The closest neighboring Face $A$ and Face $B$ for a given data utility $\mathcal{U} = 13$: 1st column) original faces; 2nd column) Standard-DP noisy faces; 3rd column) Partial-DP noisy faces; and 4th column) the optimal FaceIDP noisy faces.

### B. The Sanitized Face Images

Then, we obtained the closest neighboring face image pair according to **Definition III.6**, *i.e.*, the minimum Euclidean distance difference. For the LFW database, the obtained closest neighboring face images are shown in the 1st column of Fig. 5, which are labeled as Face $A$ and Face $B$. Next, the sanitizing noise for a given data utility $\mathcal{U} = 13$ in Eq. (25) is added to the closest neighboring face images with the 3 DP methods, *i.e.*, the Standard-DP sanitized face images in the 2nd column; the Partial-DP sanitized face images in the 3rd column; and the optimal FaceIDP sanitized face images in the 4th column. To show the difference clearly, Fig. 7 zooms in the left eye of Face $A$ and mouth of Face $B$, from which we can see that the optimal FaceIDP method obtain the most significant difference from the original face images, providing better protection for the FaceID privacy or smaller privacy budget $\epsilon$, for a given data utility $\mathcal{U}$. Similarly, for the PubFig database, Fig. 6 shows original and noisy face images of the 3 approaches, for a data

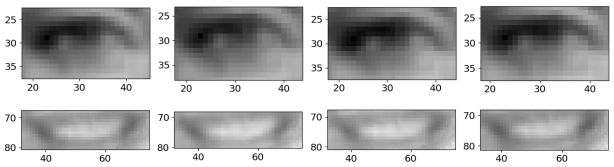

Fig. 7. (LFW Database) The zoom-in view of (left) eye and mouth of the closest neighboring Face $A$ and Face $B$ of Fig. 5, for a given data utility $\mathcal{U} = 13$: 1st column) original (left) eye and mouth; 2nd column) Standard-DP noisy (left) eye and mouth; 3rd column) Partial-DP noisy (left) eye and mouth; and 4th column) the optimal FaceIDP noisy (left) eye and mouth.

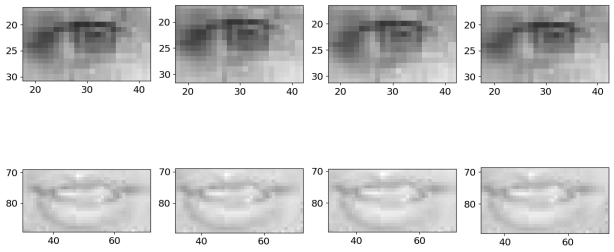

Fig. 8. (PubFig Database) The zoom-in view of (left) eye and mouth of the closest neighboring Face $A$ and Face $B$ of Fig. 5, for a given data utility $\mathcal{U} = 13$: 1st column) original (left) eye and mouth; 2nd column) Standard-DP noisy (left) eye and mouth; 3rd column) Partial-DP noisy (left) eye and mouth; and 4th column) the optimal FaceIDP noisy (left) eye and mouth.

utility $\mathcal{U} = 15$. Also, Fig. 8 shows the zoom in views of the left eye and mouth respectively, from which one can see that the privacy budget $\epsilon$ is smaller for the optimal FaceIDP method, as expected. In particular, it can be noticed that the quality of the facial images of the PubFig database (Fig. 8) is not as good as those of the LFW database (Fig. 7) due to the uncontrolled situations with non-cooperative subjects.

### C. The Sanitized Coding Coefficients and Feature Vectors

After that, to show the quantified results of the privacy protection, both the sanitized coding coefficients and the feature vectors are calculated. For the LFW database, the coding coefficients difference (dimension of 100) between Face $A$ and Face $B$ is obtained, as shown in Fig. 12. Also, the feature vector difference (dimension of 128) between Face $A$ and Face $B$ is obtained, as shown in Fig. 10, from which we can see that the sum of the feature vector difference of the optimal FaceIDP (red triangles) is the smallest for a given data utility $\mathcal{U}$, indicating better FaceID privacy protection or smaller $\epsilon$. Similarly, for the PubFig database, Fig. 9 and Fig. 11 show the the coding coefficients differences and the feature vector differences for the original face images and noisy face images of the 3 approaches respectively, from which one can see better FaceID privacy protection has been achieved for the optimal FaceIDP method.

### D. The Privacy Budget and Data Utility

To show the performance of the optimal FaceIDP, the privacy budgets $\epsilon$ for different data utilities have been obtained.

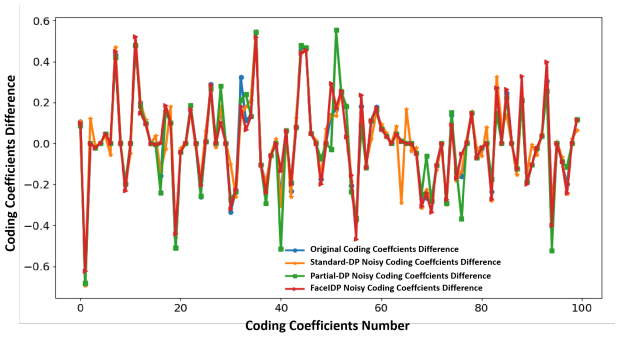


Fig. 9. (PubFig Database) The coding coefficients difference between Face $A$ and Face $B$: original value, Standard-DP noisy value, Partial-DP noisy value and the optimal FaceIDP noisy value.

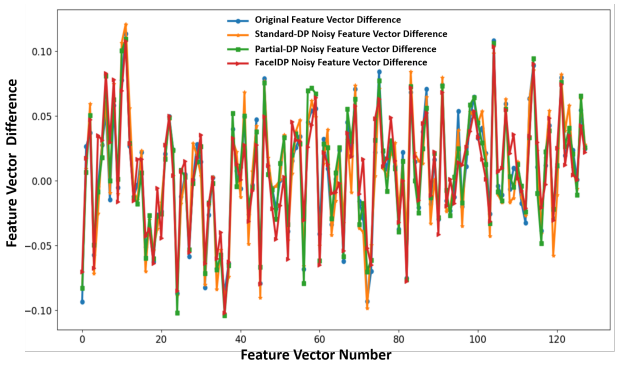


Fig. 10. (LFW Database) The feature vector difference between Face $A$ and Face $B$: original value, Standard-DP noisy value, Partial-DP noisy value and the optimal FaceIDP noisy value.

For the LFW database, $\epsilon$ is calculated for $\mathcal{U} = [5, 30]$ and the result is shown on the left plot of Fig. 13, from which it can be seen that the FaceID privacy protection of the FaceIDP method is the best among all methods, *i.e.*, it has the smallest privacy budget $\epsilon$ (green stars). Also, the Partial-DP method is better than the Standard-DP method, which is because that only the most significant coding coefficients are used in the Partial-DP method to achieve better privacy protection with smaller data utility $\mathcal{U}$. Also, on the right plot of Fig. 13, the data utility $\mathcal{U}$ is plotted against the privacy budget $\epsilon$, which again shows that the optimal FaceIDP has the smallest data utility (the best data utility) for a given privacy budget $\epsilon = (2.2, 2.9)$. Also, the Partial-DP method shows better performance than the Standard-DP method, *i.e.*, for a given privacy budget $\epsilon$, the data utility $\mathcal{U}$ is smaller (better).

Similarly, for the PubFig database, the left and right plots of Fig. 13 show the privacy budget $\epsilon$ for different data utility $\mathcal{U}$ and vice versa respectively, from which again it can be seen that the optimal FaceIDP outperforms the other 2 approaches.

## X. Conclusion

In this paper, the differential privacy problem of FaceID, *i.e.*, the FaceIDP, has been studied. Under the CS framework, a DLNet is built to learn the dictionary basis of the face images. After that, the sanitizing noise is added to the coding coefficients of the face images. Then the FaceIDP is proved to

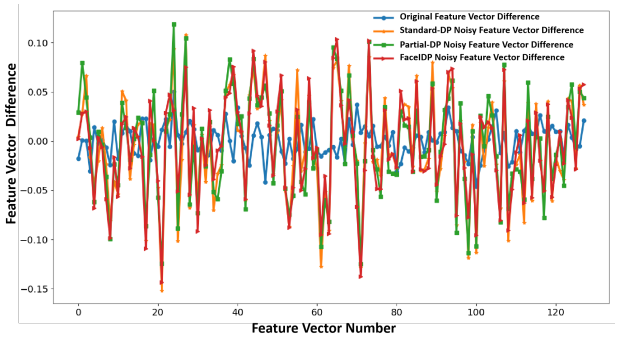


Fig. 11. (PubFig Database) The feature vector difference between Face $A$ and Face $B$: original value, Standard-DP noisy value, Partial-DP noisy value and the optimal FaceIDP noisy value.

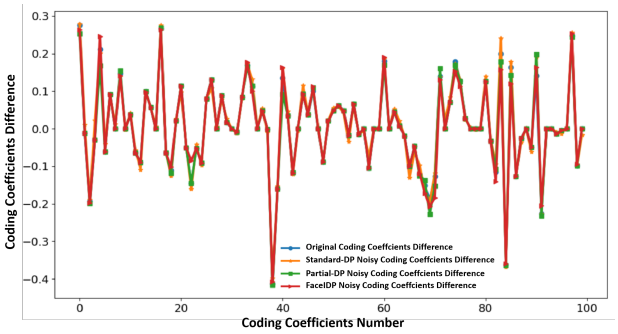


Fig. 12. (LFW Database) The coding coefficients difference between Face $A$ and Face $B$: original value, Standard-DP noisy value, Partial-DP noisy value and the optimal FaceIDP noisy value.

be $\varepsilon$-differentially private and the lower and upper bounds of the privacy budget are obtained. What's more important, the formulas of the optimal noise parameters to achieve better data utility have been derived. Also, experiment has been carried out with 2 facial images database, *i.e.*, the LFW and the PubFig databases, to confirm the efficiency of the FaceIDP to protect the FaceID privacy while still achieving good data utility. Although only 2D face identification privacy problem is studied in this paper, the FaceIDP approach can be readily extended to the 3D FaceID privacy problem. At last, the FaceIDP can be deployed in many scenarios, including face images transfer between the cloud server and the smartphones, point-to-point face images transmission, as well as face-to-face real-time video chat.

## REFERENCES


[1] 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, Jun. 2019.

[2] L. Ou, S. Liao, Z. Qin, and H. Yin, "Millimeter wave wireless hadamard image transmission for mimo enabled 5G and beyond," *IEEE Wireless Communications*, pp. 1–6, 2020.

[3] H. Wang, S. Xie, and Y. Hong, "Videodp: A flexible platform for video analytics with differential privacy," *Proceedings on Privacy Enhancing Technologies*, vol. 4, pp. 277–297, 2020.

[4] L. Ou, Z. Qin, S. Liao, T. Li, and D. Zhang, "Singular Spectrum Analysis for Local Differential Privacy of Classifications in the Smart Grid," *IEEE Internet of Things Journal*, pp. 1–1, 2020, conference Name: IEEE Internet of Things Journal.

[5] L. Ou, Z. Qin, S. Liao, Y. Hong, and X. Jia, "Releasing correlated trajectories: Towards high utility and optimal differential privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 6, pp. 1–1, 2018.

[6] L. Ou, Z. Qin, S. Liao, H. Yin, and X. Jia, "An optimal pufferfish privacy mechanism for temporally correlated trajectories," *IEEE Access*, vol. 6, pp. 37 150–37 165, 2018.

[7] A. Tonge and C. Caragea, "Image privacy prediction using deep neural networks," *ACM Transactions on the Web*, vol. 14, no. 2, pp. 7:1–7:32, Apr. 2020.

[8] S. Hill, Z. Zhou, L. Saul, and H. Shacham, "On the (in)effectiveness of mosaicing and blurring as tools for document redaction," in *Proc. Privacy Enhancing Technology*, 2016, pp. 403–417.

[9] R. McPherson, R. Shokri, and V. Shmatikov, "Defeating image obfuscation with deep learning," *arXiv*, 2016. [Online]. Available: arXiv:1609.00408

[10] M. R. Ra, R. Govindan, and A. Ortega, "P3: toward privacy-preserving photo sharing," in *Proc. the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, 2013, pp. 515–528.

[11] L. Fan, "Image pixelization with differential privacy," in *Proc. 2018 Data and Applications Security and Privacy XXXII*. Cham: Springer International Publishing, 2018, pp. 148–162.

[12] C. Tong and Z. Zheng, "An image privacy protection algorithm based on adversarial perturbation generative networks," *ACM Transactions on Multimedia Computer Communication Application*, vol. 0, no. ja, 2020. [Online]. Available: https://doi.org/10.1145/3381088

[13] J. Yang, J. Liu, and J. Wu, "Facial image privacy protection based on principal components of adversarial segmented image blocks," *IEEE Access*, vol. 8, pp. 103 385–103 394, 2020.

[14] E. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, 2008.

[15] Q. Pian, R. Yao, N. Sinsuebphon, and X. Intes, "Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging," *Nature Photonics*, vol. 11, pp. 411–414, June 2017.

[16] A. P. Spencer, B. Spokoyny, S. Ray, F. Sarvari, and E. Harel, "Mapping multidimensional electronic structure and ultrafast dynamics with single-element detection and compressive sensing," *Nature Communications*, vol. 7, no. 10434, January 2016.

[17] N. Gopalsami, S. Liao, T. W. Elmer, E. R. Koehl, A. Heifetz, A. C. Raptis, L. Spinoulas, and A. K. Katsaggelos, "Passive millimeter-wave imaging with compressive sensing," *Optical Engineering*, vol. 51, no. 9, pp. 091 614–1:9, September 2012.

[18] N. Gopalsami, T. Elmer, and S. Liao, "Compressive sampling in passive millimeter wave imaging," in *Proceedings of SPIE 8022, Passive Millimeter-Wave Imaging Technology XIV,*, vol. 80220I, 2011. [Online]. Available: DOI: 10.1117/12.886998

[19] N. Gopalsami, S. Liao, T. Elmer, A. Heifetz, and A. C. Raptis, "Compressive sampling in active and passive millimeter-wave imaging," presented at 2011 International Conference on Infrared, Millimeter, and Terahertz Waves, Houston, TX, USA, October 2-7 2011. [Online]. Available: DOI: 10.1109/irmmw-THz.2011.6105205

[20] S. D. Babacan, M. Luessi, L. Spinoulas, A. K. Katsaggelos, N. Gopalsami, T. Elmer, R. Ahern, S. Liao, and A. Raptis, "Compressive passive millimeter-wave imaging," presented at the 18th IEEE International Conference on Image Processing, Brussels, Belgium, December 29 2011. [Online]. Available: DOI: 10.1109/ICIP.2011.6116227

[21] C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature Photonics*, vol. 8, pp. 605–609, 2014.

[22] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Appl. Phys. Lett.*, vol. 93, no. 121105, 2008.

[23] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: useful tips and tricks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1527–1550, 2017.

[24] Z. Gao, L. Dai, S. Han, Z. W. C. I, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 144–153, June 2018.

[25] Z. Q. L. Ou, S. Liao and H. Yin, "Millimeter wave wireless hadamard image transmission for mimo enabled 5g and beyond," *IEEE Wireless Communications*, July 2020.

[26] S. Liao and L. Ou, "High-speed millimeter-wave 5g/6g image transmission via artificial intelligence," in *2020 Asia-Pacific*

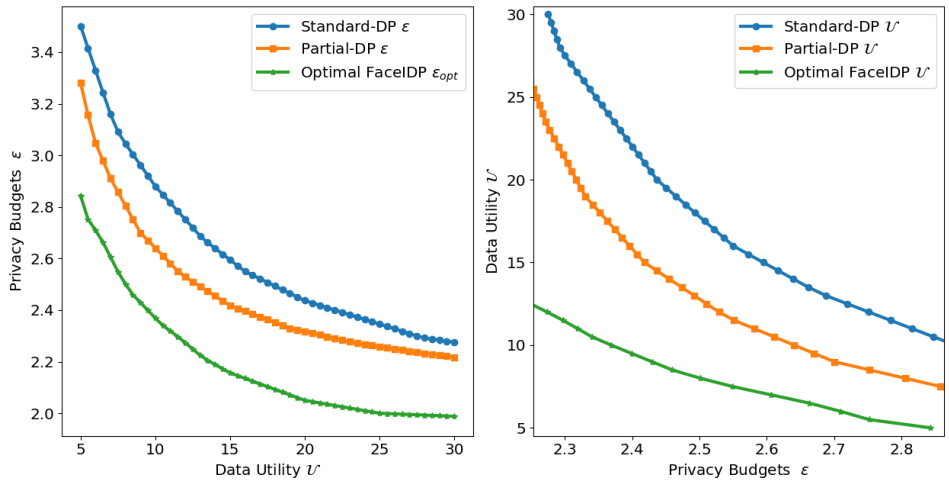


Fig. 13. (LFW Database) Privacy budget $\epsilon$ vs. data utility $\mathcal{U}$ (left) and vice versa (right).

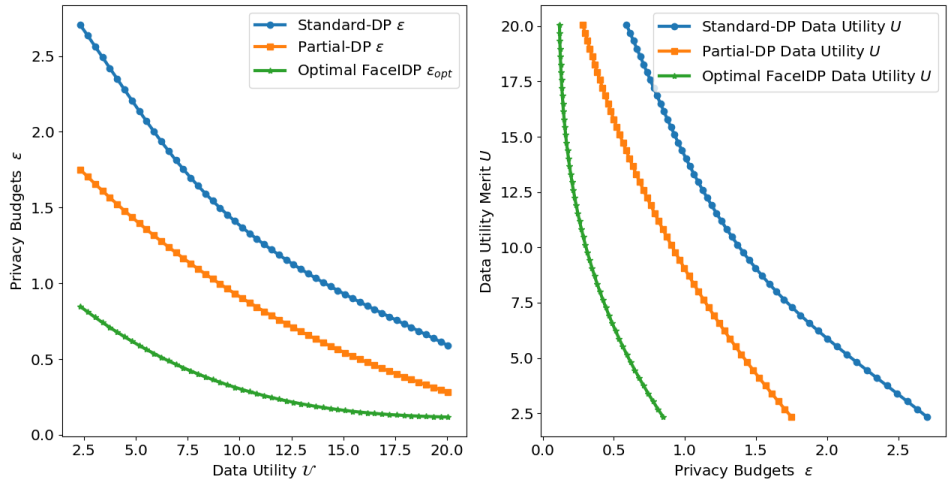


Fig. 14. (PubFig Database) Privacy budget $\epsilon$ vs. data utility $\mathcal{U}$ (left) and vice versa (right).

*Microwave Conference (APMC 2020)*, 2020. [Online]. Available: https://arxiv.org/abs/2007.03153

[27] C. Jiang, Q. Zhang, R. Fan, and Z. Hu, "Super-resolution ct image reconstruction based on dictionary learning and sparse representation," *Scientific Reports*, vol. 8, p. 8799, 2018.

[28] S. Tariyal, A. Majumdar, R. Singh, and M. Vatsa, "Deep Dictionary Learning," *IEEE Access*, vol. 4, pp. 10 096–10 109, 2016.

[29] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007.

[30] T. Murakami and Y. Kawamoto, "Utility-optimized Local Differential Privacy Mechanisms for Distribution Estimation," in *Proc. of the 28th USENIX Conference on Security Symposium*, ser. SEC'19. Berkeley, CA, USA: USENIX Association, 2019, pp. 1877–1894, event-place: Santa Clara, CA, USA.

[31] P. Kairouz, S. Oh, and P. Viswanath, "Extremal Mechanisms for Local Differential Privacy," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 492–542, Jan. 2016.

[32] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *International Conference on Computer Vision (ICCV)*, 2009.

[33] S. Liao, "Face recognition data files for dlib," IEEE Dataport, Sep. 30, 2020. [Online]. Available: http://doi.org/10.21227/r9p0-r710

**Lu Ou** received the Ph.D. degree from Hunan University in software engineering in 2018 and now is a postdoc in the College of Computer Science and Electronic Engineering at Hunan University, China.

**Zheng Qin** received the Ph.D. degree in computer software and theory from Chongqing University, China, in 2001 and now is a professor in the College of Computer Science and Electronic Engineering, Hunan University.

**Shaolin Liao** received his Ph.D. in Electrical Engineering from the University of Wisconsin, Madison, USA, in 2008 and now a Professor of Research and Adjunct Teacher at the Department of Electrical and Computer Engineering of Illinois Institute of Technology, Chicago, IL, USA.

**Yuan Hong** received his Ph.D. degree in Information Technology from Rutgers, the State University of New Jersey and now is an Assistant Professor in the Department of Computer Science at Illinois Institute of Technology.

**Dafang Zhang** received the Ph.D. degree in application mathematics from Hunan University, Changsha, China, in 1997 and now is a professor with College of Computer Science and Electronic Engineering, Hunan University, China.