SECURE FACE RECOGNITION IN EDGE AND CLOUD NETWORKS: FROM THE ENSEMBLE LEARNING PERSPECTIVE

Yitu Wang[†⋆] *Takayuki Nakachi*[†]

[†] NTT Network Innovation Laboratory, NTT Corporation, Yokosuka, Kanagawa, 239-0847 Japan

ABSTRACT

Offloading the computationally intensive workloads to the edge and cloud not only improves the quality of computation, but also creates an extra degree of diversity by collecting information from devices in service, which, in turn, has raised significant concerns on privacy as the aggregated information could be misused without the permission by the third party. Sparse coding, which has been successful in computer vision, is finding application in this new domain. In this paper, we develop a secure face recognition framework to orchestrate sparse coding in edge and cloud networks. Specifically, 1). To protect the privacy, we develop a low-complexity encrypting algorithm based on random unitary transform, where its influence on dictionary learning and sparse representation is analysed. We further prove that such influence will not affect the accuracy of face recognition. 2). To fully utilize the multi-device diversity, we extract deeper features in an intermediate space, expanded according to the dictionaries from each device, and perform classification in this new feature space to combat the noise and modeling error.

***Index Terms*—** Face Recognition, Security, Edge and Cloud, Diversity, Sparse Representation

1. INTRODUCTION

Face Recognition (FR) has been a prominent biometric technique for identity authentication [1, 2]. Significant theoretical and experimental research has been done to promote the accuracy of FR. In [3], K-Singular Value Decomposition (K-SVD) is adopted to learn a discriminative dictionary, then Orthogonal Matching Pursuit (OMP) is applied to find the sparse representation for FR. In [4], a deep convolutional neural network is utilized to extract high-level visual features, which boosts the performance of FR. However, these complex and well-engineered approaches pose exigent requirements on computing, which cannot be easily satisfied by solely relying on user devices due to their limited resources.

One promising solution is to make use of the hierarchically distributed computing structure consisting of the edge, cloud and devices [5]. In this stand, not only those computation demands can be fulfilled, but also the long latency incurred due to the information exchange in wide area networks (WAN) can be totally avoided [6]. In [7, 8], the computational efficiency of FR is improved by distributing a part of the computation tasks to the edge and cloud. However, such strategies suffer from one major drawback, i.e., the extra degree of diversity generated through aggregating information from multiple devices is neglected.

To exploit more dimensions of the edge and cloud resources, we construct a framework to not only reduce the computation demands at each device, but also take the advantage of the multi-device diversity to produce a more accurate FR result. The motivation and main contributions of this paper are summarized as follows,

1. **Preserve the privacy by random unitary transform:** Encrypting algorithms allowing computation on ciphertexts, such as Homomorphic Encryption (HE) and secure Multi-Party Computation (MPC) [9], are faced with the curse of dimensionality regarding the size of images. To address this challenge, we develop a low-complexity encrypting algorithm based on random unitary transform, which enables that dictionaries/FR results can be directly trained/drawn from the encrypted images. Moreover, we theoretically prove that such encryption will not affect the accuracy of FR.

2. **Exploit multi-device diversity by ensemble learning:** The performance of the sparse coding-based FR algorithms rely heavily on the number of training samples, while the excessive cost of bandwidth and computation makes it difficult to gather all the training samples for dictionary learning. Alternatively, we integrate only the dictionaries from each device, based on which we obtain the decision templates for each class, in order to extract the combined effect of noise and modeling error. Given a testing image, the FR result is obtained according to the pairwise similarity between its decision profile and each of the decision templates to reduce the influence of noise and modeling error on FR.

The rest of this paper is organized as follows. Section 2 presents the system model. In Section 3, we propose the secure FR framework. Following this, the performance of the proposed framework is evaluated in Section 4 by simulation. Finally, this paper concludes with Section 5.

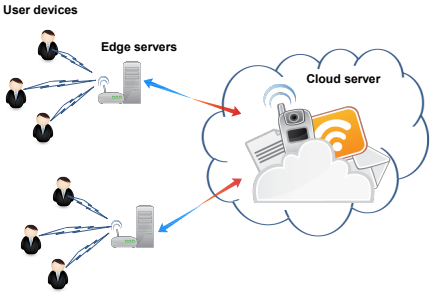


Fig. 1. The Architecture of the Edge and Cloud Networks

2. SYSTEM MODEL

2.1. The Edge and Cloud Networks

We first introduce the communication model for the edge and cloud networks as shown in Fig. 1, where N single core user devices, denoted as set $\mathcal{N}$, are assisted by M edge servers and one remote cloud. Among the devices, L classes of individuals are to be recognized, denoted as $\mathcal{L}$, and the training set of class i and device j is denoted as $\boldsymbol{B}_i^j$ [10]. Each edge server is a light-weight computing center deployed at a wireless access point, while the remote cloud connects with each edge server using the backbone network [11]. A device will offload its computation tasks to the edge server in close proximity via a wireless channels[1], the edge server together with the cloud will execute the computation tasks on behalf of the device.

2.2. Sparse Representation based FR

To classify according to the sparse representation of face images, we proceed by two steps:

1) Dictionary Training:

Given an m-dimensional training set $\boldsymbol{B}^j \in \mathbb{R}^{m\times|\boldsymbol{B}^j|}$ of device $j \in \mathcal{N}$, we jointly train a dictionary with K^j atoms and classifier parameters for FR as in [12],

$$\begin{aligned} P(\boldsymbol{B}^j) = \arg \min_{\boldsymbol{X}^j,\boldsymbol{D}^j,\boldsymbol{W}^j,\boldsymbol{A}^j} ||\boldsymbol{Z}^j - \boldsymbol{T}^j\boldsymbol{X}^j||_2^2 \\ s.t. ||\boldsymbol{x}_i^j||_0 \leq \epsilon, \forall i \in \{1,2,\cdots,|\boldsymbol{B}^j|\}, \end{aligned} \tag{1}$$

where

$$\boldsymbol{Z}^j = \begin{bmatrix} \boldsymbol{B}^j \\ \sqrt{\alpha}\boldsymbol{C}^j \\ \sqrt{\beta}\boldsymbol{H}^j \end{bmatrix}, \boldsymbol{T}^j = \begin{bmatrix} \boldsymbol{D}^j \\ \sqrt{\alpha}\boldsymbol{A}^j \\ \sqrt{\beta}\boldsymbol{W}^j \end{bmatrix}, \tag{2}$$

in which $\boldsymbol{D}^j \in \mathbb{R}^{m\times K^j}$ represents the dictionary, $\boldsymbol{X}^j \in \mathbb{R}^{K^j\times|\boldsymbol{B}^j|}$ denotes the sparse representation, α and β are the weights for the label consistent term and the reconstruction error term, $\boldsymbol{C}^j$ is the discriminative sparse code, $\boldsymbol{A}^j$ is the linear transformation matrix, $\boldsymbol{W}^j$ is the classifier matrix, $\boldsymbol{H}^j$ is the class label matrix, and ϵ is the sparsity constraint. This problem can be solved efficiently using K-SVD [3].

[1]Some physical layer access scheme, e.g., Code Division Multiple Access (CDMA), is adopted to allow multiple devices to share the same edge server simultaneously. The matching between devices and edge servers is accomplished during network setup period [11].

2) Face Recognition:

Given a testing sample $\boldsymbol{y}^k, k \in \mathcal{N}$, the sparse representation based on $\boldsymbol{D}^j$ can be calculated according to

$$\arg\min_{\boldsymbol{x}_j^k} ||\boldsymbol{y}^k - \boldsymbol{D}^j\boldsymbol{x}_j^k||_2^2 \; s.t. \; ||\boldsymbol{x}_j^k||_0 \leq \epsilon, \tag{3}$$

which can be solved efficiently using OMP [3]. Then, the class label for this testing sample $\boldsymbol{y}^k$ can be estimated as $l_j^k = \arg\max\{\boldsymbol{W}^j\boldsymbol{x}_j^k\}$. Although such a scheme has been proved to be effective, it faces two major drawbacks,

1). FR is performed based on the local dictionary only, i.e., $k = j$, which makes it vulnerable to noise and modeling error.

2). The inherent information of the class label vector $\boldsymbol{W}\boldsymbol{x}$ is not fully utilized due to the *argmax* operation.

2.3. Problem Formulation

To address the above two problems, we extend the conventional FR algorithm, which is further applied to the edge and cloud networks. In this paper, our objective is to

1). Protect the privacy regarding the information passing during the whole process.

2). Propose a communicationally efficient algorithm to exploit the multi-device diversity and produce a refined FR result.

3. SECURE FR IN EDGE AND CLOUD NETWORKS

3.1. Random Unitary Transform

Random unitary transform not only proves to be effective for biometric template protection, but also with desired low computational complexity [13]. Any vector $\boldsymbol{v} \in \mathbb{R}^{m\times 1}$ encrypted by random unitary matrix $\boldsymbol{Q}_p \in \mathbb{C}^{m\times m}$ with a private key p can be expressed as follows,

$$\overline{\boldsymbol{v}} = f(p,\boldsymbol{v}) = \boldsymbol{Q}_p\boldsymbol{v}, \tag{4}$$

where $\overline{\boldsymbol{v}}$ is the encrypted vector, and $\boldsymbol{Q}_p$ satisfies

$$\boldsymbol{Q}_p^*\boldsymbol{Q}_p = \boldsymbol{I}, \tag{5}$$

where $[\cdot]^*$ and $\boldsymbol{I}$ represents the Hermitian transpose and identity matrix, respectively. Gram-Schmidt orthogonalization can be adopted for generating $\boldsymbol{Q}_p$. The encrypted vector has three properties [14] as follows,

1). *Conservation of the Euclidean distances*: $||\boldsymbol{v}_i - \boldsymbol{v}_j||_2^2 = ||\overline{\boldsymbol{v}}_i - \overline{\boldsymbol{v}}_j||_2^2$,

2). *Norm isometry*: $||\boldsymbol{v}||_2^2 = ||\overline{\boldsymbol{v}}||_2^2$,

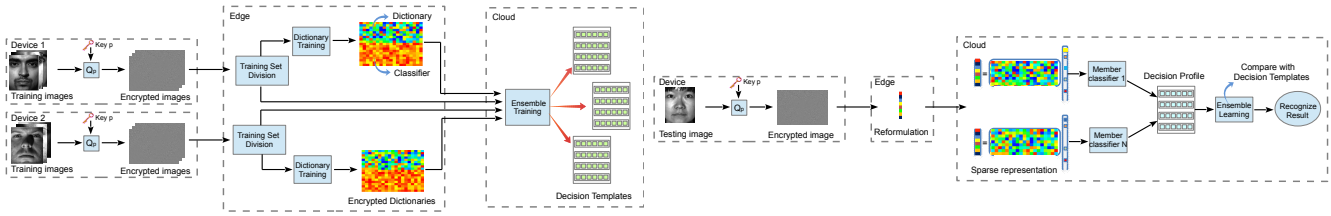
3). *Conservation of inner products*: $\boldsymbol{v}_i \times \boldsymbol{v}_j^T = \overline{\boldsymbol{v}}_i \times \overline{\boldsymbol{v}}_j^T$.

3.2. Secure FR

According to random unitary transform, the encrypted training samples $\overline{\boldsymbol{B}}^j$ and testing samples $\overline{\boldsymbol{Y}}^k$ are generated as

$$\overline{\boldsymbol{B}}^j = f(p,\boldsymbol{B}^j) = \boldsymbol{Q}_p\boldsymbol{B}^j, \; \overline{\boldsymbol{Y}}^k = f(p,\boldsymbol{Y}^k) = \boldsymbol{Q}_p\boldsymbol{Y}^k. \tag{6}$$

To obtain the dictionary and classifier matrix based on the encrypted training samples, we should solve Eq. (1) using $\overline{\boldsymbol{B}}^j$ instead of $\boldsymbol{B}^j$. In the following theorem, we demonstrate the influence of random unitary transform on the trained dictionary and the classifier matrix.



(a) Ensemble Training (b) Recognizing

Fig. 2. The Proposed Framework

Theorem 1. *The trained dictionary* $\overline{\boldsymbol{D}}^j$ *and the classifier matrix* $\overline{\boldsymbol{W}}^j$ *based on the encrypted training samples* $\overline{\boldsymbol{B}}^j$ *satisfy*

$$\overline{\boldsymbol{D}}^j = \boldsymbol{Q}_p \boldsymbol{D}^j,\ \overline{\boldsymbol{W}}^j = \boldsymbol{W}^j, \tag{7}$$

where $\boldsymbol{D}^j, \boldsymbol{W}^j$ *are those trained from the unencrypted training samples.* ■

Proof. Solving Eq. (1) involves iterating the two steps below, 1) Sparse coding step: Given the dictionary $\boldsymbol{T}^j$, the objective is to obtain the best sparse coefficient $\boldsymbol{X}^j$. We apply OMP to solve this problem, please refer to [16] for the details.

First, in the *Sweep* step, we compute the error $\overline{\epsilon}^j(i)$, where i denotes the label of atom in $\boldsymbol{T}^j$, as

$$\begin{aligned}\overline{\epsilon}^j(i) =& ||\boldsymbol{Q}_p\boldsymbol{B}^j||_2^2 + ||\sqrt{\alpha}\boldsymbol{C}^j||_2^2 + ||\sqrt{\beta}\boldsymbol{H}^j||_2^2 \\ &- \frac{(\boldsymbol{Q}_p\boldsymbol{B}^j\boldsymbol{Q}_p\boldsymbol{D}_i^j + \alpha\boldsymbol{C}^j\boldsymbol{A}_i^j + \beta\boldsymbol{H}^j\boldsymbol{W}_i^j)^2}{||\boldsymbol{Q}_p\boldsymbol{D}^j||_2^2 + ||\sqrt{\alpha}\boldsymbol{A}^|||_2^2 + ||\sqrt{\beta}\boldsymbol{W}^j||_2^2}.\end{aligned} \tag{8}$$

According to the three properties, we have $\overline{\epsilon}^j(i) = \epsilon^j(i)$. Hence, the minimizer is not influenced, which accounts for the same support S_i.

Second, in the *Update Provisional Solution* step, we update the sparse representation according to the i-th atom with the minimum error, as

$$\begin{aligned}\overline{E}_i^j =& ||\boldsymbol{Q}_p\boldsymbol{B}^j - \boldsymbol{Q}_p\boldsymbol{D}_{S_i}^j\boldsymbol{X}_{S_i}^j||_2^2 + ||\sqrt{\alpha}\boldsymbol{C}^j - \sqrt{\alpha}\boldsymbol{A}_{S_i}^j \\ &\times \boldsymbol{X}_{S_i}^j||_2^2 + ||\sqrt{\beta}\boldsymbol{H}^j - \sqrt{\beta}\boldsymbol{W}_{S_i}^j\boldsymbol{X}_{S_i}^j||_2^2.\end{aligned} \tag{9}$$

According to property norm isometry, we have $\overline{E}_i^j = E_i^j$. Therefore, the provisional solution is not affected.

Finally, in the *Stopping Rule* step, similarly, we have $||\overline{r}_i^j||_2^2 = ||r_i^j||_2^2 \leq \epsilon$. Therefore, the above analysis proves that random unitary transform does not affect the sparse coding step under the condition $\overline{\boldsymbol{D}}^j = \boldsymbol{Q}_p\boldsymbol{D}^j$.

2) Dictionary update step: Given the sparse coefficient $\boldsymbol{X}^j$, the objective is to find the dictionary that best describes the training samples. We apply K-SVD to find the solution [17].

In the *Compute the residual matrix* step, the representation error matrix $\boldsymbol{E}_d^j$ for the d-th column is

$$\boldsymbol{E}_d^j = \boldsymbol{Z}^j - \sum_{i\neq d}^{K^j} \boldsymbol{t}_i^j(\boldsymbol{x}_i^j)^T, \tag{10}$$

where $\boldsymbol{t}_i^j$ represents the i-th column in $\boldsymbol{T}^j$, $(\boldsymbol{x}_i^j)^T$ represents the i-th row in $\boldsymbol{X}^j$. To minimize the l^2-norm of $\boldsymbol{E}_d^j$ while keeping the cardinalities of all the representations fixed, we restrict $\boldsymbol{E}_d^j$, by choosing only the columns where the entries in the row are non-zero, into $(\boldsymbol{E}_d^j)^R$. Then apply SVD,

$$\begin{aligned}(\boldsymbol{E}_d^j)^R &= \begin{bmatrix}(\boldsymbol{B}^j - \sum_{i\neq d}^{K^j}\boldsymbol{D}_i^j(\boldsymbol{x}_i^j)^T)^R \\ (\sqrt{\alpha}\boldsymbol{C}^j - \sum_{i\neq d}^{K^j}\sqrt{\alpha}\boldsymbol{A}_i^j(\boldsymbol{x}_i^j)^T)^R \\ (\sqrt{\beta}\boldsymbol{H}^j - \sum_{i\neq d}^{K^j}\sqrt{\beta}\boldsymbol{W}_i^j(\boldsymbol{x}_i^j)^T)^R\end{bmatrix} \\ &= \begin{bmatrix}\boldsymbol{U}_A^j & 0 \\ 0 & \boldsymbol{U}_B^j\end{bmatrix}\begin{bmatrix}\boldsymbol{S}_A^j & 0 \\ 0 & \boldsymbol{S}_B^j\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_A^j \\ \boldsymbol{V}_B^j\end{bmatrix} = \sum_{i=1}^{m}\boldsymbol{u}_i^j\cdot\boldsymbol{\sigma}_i^j(\boldsymbol{v}_i^j)^T.\end{aligned} \tag{11}$$

As for the representation error matrix $\overline{\boldsymbol{E}}_d^j$ using the encrypted data, we have $(\boldsymbol{Q}_p\boldsymbol{B}^j - \sum_{i\neq d}^{K^j}\overline{\boldsymbol{D}}_i^j(\boldsymbol{x}_i^j)^T)^R = \boldsymbol{Q}_p\boldsymbol{U}_A^j\boldsymbol{S}_A^j(\boldsymbol{V}_A^j)^T$, derived from the scaling property of the optimal minimizer, and $[(\sqrt{\alpha}\boldsymbol{C}^j - \sum_{i\neq d}^{K^j}\sqrt{\alpha}\overline{\boldsymbol{A}}_i^j(\boldsymbol{x}_i^j)^T)^R; (\sqrt{\beta}\boldsymbol{H}^j - \sum_{i\neq d}^{K^j}\sqrt{\beta}\overline{\boldsymbol{W}}_i^j(\boldsymbol{x}_i^j)^T)^R] = \boldsymbol{U}_B^j\boldsymbol{S}_B^j(\boldsymbol{V}_B^j)^T$, which is obtained according to the invariant property of the optimal minimizer when enforcing $\tilde{\boldsymbol{x}}_i^j = \boldsymbol{x}_i^j$.

Then, we apply SVD to minimize the l^2-norm of $\overline{\boldsymbol{E}}_d^j$, and update the dictionary, classifier and sparse coefficient as

$$\begin{aligned}\overline{\boldsymbol{D}}_d^j &= \overline{\boldsymbol{u}}_1^j = \boldsymbol{Q}_p\boldsymbol{u}_1^j = \boldsymbol{Q}_p\boldsymbol{D}_d^j \\ <\overline{\boldsymbol{A}}_d^j, \overline{\boldsymbol{W}}_d^j> &= \overline{\boldsymbol{u}}_{Kj+1}^j = \boldsymbol{u}_{Kj+1}^j =< \boldsymbol{A}_d^j, \boldsymbol{W}_d^j > \\ \tilde{\boldsymbol{x}}_d^j &= \overline{\boldsymbol{\sigma}}_1^j(\overline{\boldsymbol{v}}_1^j)^T = \boldsymbol{\sigma}_1^j(\boldsymbol{v}_1^j)^T = \boldsymbol{x}_d^j.\end{aligned} \tag{12}$$

□

To classify the testing sample $\overline{\boldsymbol{Y}}^k$, we obtain its sparse representation based on $\overline{\boldsymbol{D}}^j$ by solving

$$<\tilde{\boldsymbol{X}}_k^j>= \arg\min_{\boldsymbol{x}_k^j} ||\overline{\boldsymbol{Y}}^k - \overline{\boldsymbol{D}}^j\boldsymbol{X}_k^j||_2^2 \ \ s.t. ||\boldsymbol{x}_k^j||_0 \leq \epsilon. \tag{13}$$

It is proved in the *Sparse coding step* that the sparse representation will not be affected by random unitary transform.

Remark 1. *According to the above analysis, the classifier* $\overline{\boldsymbol{W}}^j$ *and sparse representation* $\tilde{\boldsymbol{X}}_k^j$ *trained/drawn from the encrypted training/testing images are identical to those trained/drawn from the unencrypted ones, which renders the class label vectors identical. Therefore, the privacy can be preserved without performance degradation.* ■

3.3. Ensemble Learning Framework

We introduce decision profile and decision template for better utilizing the class label vector in the edge and cloud network.

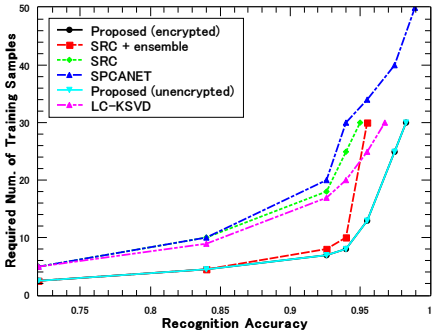


Fig. 3. Performance Comparison (10 devices are deployed)

1). **Decision Profile (DP)**: A DP is a matrix with size $L \times N$, where the j-th column represents the normalized class label vector based on $\overline{\boldsymbol{D}}^j$. To better exploit the information within the class label vectors in a holistic manner, we treat the values as features in the *intermediate feature space*.

2). **Decision Template (DT)**: A DT is to remember the most typical DP for each class and extract the combined effect of the noise and modeling error. The classification result will be carried out by measuring the pairwise similarity between the DP of the testing sample and each of the DTs.

The proposed framework mainly consists of two stages,

1). **Ensemble Training:**

i). *Training Set Division*: We partition the received training samples $\boldsymbol{B}^j$ into two parts at the edge servers, one for dictionary training, i.e., $\boldsymbol{B}^{j,D}$, and the other for ensemble training.

ii). *Dictionary and Classifier Training*: We jointly train the dictionary $\boldsymbol{D}^j$ and classifier $\boldsymbol{W}^j$ based on $\boldsymbol{B}^{j,D}$.

iii). *Ensemble Training*: The ensemble training set can be formulated accordingly as $\boldsymbol{B}^E$ at the cloud. Then, we estimate the class label vector for $\boldsymbol{y} \in \boldsymbol{B}^E$ using its sparse representation $\boldsymbol{x}^j$ under $\boldsymbol{D}^j$, and the predictive classifier $\boldsymbol{W}^j$ as

$$\boldsymbol{L}(\boldsymbol{x}^j) = \{\boldsymbol{W}^j \boldsymbol{x}^j\}, \forall j \in \mathcal{N}. \tag{14}$$

Next, we formulate its DP as

$$\boldsymbol{DP}(\boldsymbol{y}) = [\boldsymbol{L}(\boldsymbol{x}^1), \boldsymbol{L}(\boldsymbol{x}^2), \cdots, \boldsymbol{L}(\boldsymbol{x}^N)]. \tag{15}$$

Finally, the DT for class $i \in \mathcal{L}$ can be calculated as

$$\boldsymbol{DT}_i = \frac{1}{|\boldsymbol{B}_i^E|} \sum_{\boldsymbol{y} \in \boldsymbol{B}_i^E} \boldsymbol{DP}(\boldsymbol{y}). \tag{16}$$

2). **Recognizing:**

First, upon receiving a testing image $\boldsymbol{y}^k$, the edge server only transmits the feature descriptor extracted by random faces [3] to the cloud. Then, we obtain its $\boldsymbol{DP}(\boldsymbol{y}^k)$. Finally, we measure its pairwise similarity against $\boldsymbol{DT}_i, \forall i \in \mathcal{L}$ as

$$\mu(\boldsymbol{DP}(\boldsymbol{y}^k), \boldsymbol{DT}_i) = \frac{1}{L \times N} \sum_{l \in \mathcal{L}} \sum_{j \in \mathcal{N}} \left(\boldsymbol{DT}_i(l,j) - \boldsymbol{DP}_{l,j}(\boldsymbol{y}^k)\right)^2. \tag{17}$$

The classification is identified as the class label i^*, where $\mu(\boldsymbol{DP}(\boldsymbol{y}^k), \boldsymbol{DT}_{i^*}) < \mu(\boldsymbol{DP}(\boldsymbol{y}^k), \boldsymbol{DT}_i), \forall i \neq i^*$.

Remark 2. *The training samples are differently and independently chosen according to devices, and such uniqueness of the information available in each training set prompts the dictionaries to capture different patterns, which accounts for the multi-device diversity. Moreover, we try to estimate the expectation of the noise and modeling error based on $\boldsymbol{B}^E$, which is further included in the DTs. The similarity is calculated in a pairwise manner to reduce such influence.*

Table 1. Execution Time

Algorithm	Training Time (s)	Testing Time (s)
Proposed	7.29	1.64×10^{-3}
SPCANET	5780	1.20
LC-KSVD	4.84	1×10^{-4}
SRC	0.30	0.22
SRC+ensemble	0.30	1.11

4. SIMULATION RESULTS

We compare the performance of the proposed framework with the following four base-line algorithms,

1). *Baseline 1 (SPCANET) [4]*: A 5-layer Convolutional Neural Network (CNN) is adopted to extract more discriminative features, especially the non-linear features.

2). *Baseline 2 (SRC) [10]*: Face subspace model is adopted for sparse representation-based FR.

3). *Baseline 3 (LC-KSVD) [12]*: The label consistent term is introduced to improve the discriminative capability.

4). *Baseline 4 (SRC + ensemble)*: SRC is combined with ensemble learning in the edge and cloud network.

The Extended YaleB database is adopted for FR [19, 20]. We randomly select 32 images for each individual as the dictionary training set, 10 images for each individual as the ensemble training set, where half of them overlap with those in the dictionary training set, while the rest for testing.

In Fig. 3, first, the proposed framework outperforms all the sparse-representation based algorithms, because we take the full advantage of the multi-device diversity. Second, by adopting random unitary transform, the result of FR is not influenced. Finally, compared with SPCANET, the proposed framework achieves the same performance with less number of training samples. Especially, when there are only 10 training samples/class, which is reasonable due to the scarcity of manually labeled data, the proposed framework outperforms SPCANET by over 10%. Moreover, in Table. 1, the proposed framework is computationally efficient, which makes it possible to support secured and real-time FR applications.

5. CONCLUSIONS

We develop a secured framework for FR in the edge and cloud networks. To guarantee the privacy, random unitary transform is adopted to enable computing on cipher-texts without influencing the accuracy of FR. To exploit the multi-device diversity, 1). We obtain the DT for each class to extract the combined effect of noise and modeling error. 2). The recognition is identified according to the pairwise similarity to eliminate such influence for a refined FR result.

6. REFERENCES

[1] Y. Duan, J. Lu, J. Feng, and J. Zhou, "Context-aware local binary feature learning for face recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 40, no. 5, pp. 1139-1153, May 2018.

[2] X. Fontaine, R. Achanta, and S. Susstrunk, "Face recognition in real-world images," *Proc. of IEEE ICASSP 2017*, pp. 1482-1486, Mar. 2017.

[3] Q. Zhang, and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," *Proc. of IEEE CVPR 2010*, pp. 2691-2698, Jun. 2010.

[4] L. Tian, C. Fan, Y. Ming, and Y. Jin, "Stacked PCA network (SPCANET): an effective deep learning for face recognition," *Proc. of IEEE ICDSP 2015*, pp. 1039-1043, Jul. 2015.

[5] Y. Mao, J. Zhang, S. Song, and K. Letaief, "Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5994-6009, Jun. 2017.

[6] S. Teerapittayanon, B. McDanel, and H. T. Kung, "Distributed deep neural networks over the cloud, the edge and end devices," *Proc. of IEEE ICDCS 2017*, pp. 328-339, Jun. 2017.

[7] P. Hu, H. Ning, T. Qiu, H. Song, Y. Wang, and X. Yao, "Security and privacy preservation scheme of face identification and resolution framework using fog computing in internet of things," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1143-1155, Oct. 2017.

[8] P. Hu, H. Ning, T. Qiu, Y. Zhang, and X. Luo, "Fog computing based face identification and resolution scheme in internet of things," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1910-1920, Aug. 2017.

[9] M. Dias, A. Abad, and I. Trancoso, "Exploring hashing and cryptonet based approaches for privacy-preserving speech emotion recognition," *Proc. of IEEE ICASSP 2018*, pp. 2057-2061, Apr. 2018.

[10] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 31, no. 2, pp. 210-227, Feb. 2009.

[11] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2795-2808, Oct. 2016.

[12] Z. Jiang, Z. Lin, and L. Davis, "Learning a discriminative dictionary for sparse coding via label consistent K-SVD," *Proc. of IEEE CVPR 2011*, pp. 1697-1704, Jun. 2011.

[13] T. Nakachi, H. Ishihara, and H. Kiya, "Privacy-preserving network BMI decoding of covert spatial attention," *Proc. of IEEE ICSPCS 2018*, pp. 1-8, Dec. 2018.

[14] T. Maekawa, T. Nakachi, S. Shiota, and H. Kiya, "Privacy-preserving SVM computing by using random unitary transformation," arXiv:1809.07055, Sept. 2018.

[15] Y. Saito, I. Nakamura, S. Shiota, and H. Kiya, "An efficient random unitary matrix for biometric template protection," *Joint Proc. of IEEE SCIS 2016 and ISIS 2016,* pp. 366-370, Aug. 2016.

[16] T. Nakachi, and H. Kiya, "Practical secure OMP computation and its application to image modeling," *Proc. of ACM ICIHIP 2018*, pp. 25-29, Sept. 2018.

[17] T. Nakachi, Y. Bandoh and H. Kiya, "Secure dictionary learning for sparse representation," *Proc. of EURASIP EUSIPCO 2019*, pp. 1-5, Sept. 2019.

[18] L. Yu, S. Wang, and K. Lai, "Credit risk assessment with a multistage neural network ensemble learning approach," *Expert syst. appl.*, vol. 34, no. 2, pp. 1434-1444, Feb. 2008.

[19] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 6, pp. 643-660, Jun. 2001.

[20] K. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 5, pp. 684-698, May 2005.