Person Identity Verification Using Wi-Fi Based Handwritten Signature Signals on a Triplet Network

Young-Woong Kwon

The Graduate School
Yonsei University
School of Electrical and Electronic Engineering

Person Identity Verification Using Wi-Fi Based Handwritten Signature Signals on a Triplet Network

A Masters Thesis
Submitted to the School of Electrical and Electronic Engineering
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical and Electronic Engineering

Young-Woong Kwon

October 2019

This certifies that the masters thesis of
Young-Woong Kwon is approved.

Thesis Supervisor: Prof. Kar-Ann Toh

Thesis Committee Member: Prof. Kwanghoon Sohn

Thesis Committee Member: Prof. Andrew Beng Jin Teoh

The Graduate School

Yonsei University

September 2019

ACKNOWLEDGEMENT

October 2019

Young-Woong Kwon

Contents

List of Figures

List of Tables

Abstract

An In-Air Signature Identification System using Commercial Wi-Fi Devices

Young-Woong Kwon

School of Electrical and Electronic Engineering

The Graduate School

Yonsei University

Identity verification using Wi-Fi in-air handwritten signature is a challenge task in signature verification since the shape of the signal varies according to the direction in which the signature is written. By using the handcraft features, it was difficult to verify identity from signatures entered in various directions. Moreover, the limited size of dataset also limited the training of deep learning models. In this paper, we propose a system for identity verification from Wi-Fi in-air signature signals based on triplet network. Three-channel ConvNet structures is adopted in order to learn discriminative features from the in-air signatures. Moreover, we propose a input triplet mining approach based on the kernel and range space learning to faster the convergence speed. Our experimental results on in-house Wi-Fi handwritten signature dataset shows the proposed network outperforms

handcraft methods and Siamese network by improved verification accuracy and faster loss convergence.

Key words : Biometrics, In-air handwritten signature verification, Wi-Fi Channel State Information, triplet network and the Kernel and Range space learning

Chapter 1

Introduction

1.1 Background

Signatures written on paper have long been used as an identification tool, and various methods for verifying individuals have been developed [1–4]. As mobile devices increase, identity verification using biometrics traits is becoming more widely used. Compared to the use of complex passwords, using biometric systems to user verification has advantages of increased security, convenience, and accountability [5]. With the development of various sensors such as depth camera and mobile camera, in-air signature has also become available [6]. However, they had to use a special sensor prepared for signature. [7, 8] used depth camera, [9] used mobile device to record position of hand. On the other hand, some studies use Wi-Fi CSI signals to identify the biometric characteristics of the human body. Since they use commercial devices that are already widespread, they do not need any special input devices. [10–12] studied CSI signals to identify the various biometrics characteristics of the body. [13, 14] used CSI signals to enable users to recognize their gestures. Recently, a study was conducted to identify signatures written in the air using CSI signals [15]. However, in this study, only the signals entered in one direction were recognized. Due to the nature of the in-air signature, which is difficult to specify the direction in which the signal is input, a verification system is required regardless of the direction in which the signal is entered. More recently, there have been studies using deep learning technology to characterize CSI signals. Deep Learning-based models are spotlighted for their automated feature extractors and superior classification capabilities based on them, compared to traditional handcraft models. Deep learning was used to recognize users based on the body shape [16], or to identify them with the characteristics shown in their behavior [17]. We used deep learning technology to create a system that can be identity-recognizable even for multi-way air signatures entered with Wi-Fi CSI. The deep learning model used a triplet network to increase the accuracy of feature extraction while allowing accurate classification, and also improved the model's convergence speed using the kernel and range space running [18].

1.2 Motivation and Contributions

The main contributions of our work can be summarized as follows:

- A system to utilize Wi-Fi handwritten signature signals for identity verification is proposed. We used the deep triplet network to obtain the discriminative features in limited training dataset.

- In training of the triplet network, we propose hard triplet mining strategy using the kernel and range (KAR) space learning to faster the convergence speed.

- We propose an experimental study using an in-house Wi-Fi handwritten signature dataset collected from 98 subjects.

1.3 Organization of Thesis

The rest of this paper is organized as follows: Section 2 introduces the triplet network and the Kernel and Range space learning. Our proposed system is described in Section 3. In Section 4, experimental results is discussed. Conclusions and future works are presented in Section 5.

Chapter 2

Preliminaries

2.1 Wi-FI Channel State Information

CSI captures signal strength and phase information for OFDM subcarriers and between each pair of transmit-receive antennas. It runs on a commodity 802.11n NIC and records Channel State Information (CSI based on the 802.11 standards. The CSI contains information about the channel between sender and receiver at the level of individual data subcarriers, for each pair of transmitting and receive antennas. In a frequency domain, the CSI of sub-carrier **c** between transmitter(Tx) and receiver(Rx) can be modeled as $R_c = \mathbf{H}_c T_c + N$ where the R_c and T_c denote the received and the transmitted signal vector of dimension r and t, respectively. The N is the additive channel noise and $\mathbf{H}_c$ is the $r \times t$ channel matrix. The CSI of sub-carrier c can be modeled as follows:

$$h_c = \mid h_c \mid e^{\angle\theta}, \tag{2.1}$$

where $\mid h_c \mid$ and θ represent the amplitude and the phase of the sub-carrier, respectively.

Chapter 3

Related Works

3.1 Deep Metric Learning

Metric learning aims to learn a distance function that places similar data close together in a feature space and dissimilar data away from each other. By metric learning, multi-dimensional input data is converted to low-level feature vectors to measure the distance between data for a specific task. Metric running has been used for computer vision tasks such as image classification and content-based image retrieval [19]. Previously, features are extracted using HOG, LBP, and so on. Simailarity measured based on these features. In recent years, deep learning-based methods have been widely used, as they have enabled feature extraction and metric running in one framework [20]. The features for image classification are automatically learned by the deep learning network. Deep Learning-based models include the Siamise network and the Triplet network, which are described below.

3.1.1 Triplet network

Triplet network is also a metric running model [21] which receives triplet data as its inputs and aims to learn distance metric of the data in feature space [19]. Triplet network receives triplet data pairs as input. Triplet network is widely used for person re-identification tasks which aims to identify individuals in several images. Person re-identification handles similar objectives with the biometric identification task, but it is more challanging due to the low quality and high variety of its target images [20]. Since person re-identification task usually deals with images taken from low-resolution devices such as surveilance cameras. Also, different occasions such as the clothes, poses, and angle of the target person makes it difficult to identify.

Input triplet data is composed of anchor(reference), positive(similar) and negative(dissimilar) samples. The training of the triplet network is making feature vectors to be placed in the appropriate feature space, making the positive(similar) data is close to the anchor(reference) and the negative(dissimilar) data is kept away from the anchor. Since myriad triplet pairs may exist for the training set, training from all possible pairs may time-consuming and unnecessary. Optimizing training becomes necessary as it mines the learning-efficient triplet out of large possible inputs. [22–24] generates triplet only for small number of classes, which randomly selected in iteration. Rescently, in [25] used triplet mining stratege for faster convergence speed. They selects inputs from large mini-batch at each training iteration using the network during training. However, to make such a large-sized mini-batch, a lot of training data is needed for every iteration. Real-world data obtained from in-house experiments is not suitable for this training stratege since it is usually small in size and consists of dozens of classes. In order to optimize input triplet data with relatively small sized data, we adopt the kernel and range space manipulation method for the training.

3.1.2 Siamese networks

A Siamese neural networks consists of twin networks which accept distinct inputs but are joined by an energy function at the top [26]. By using a constrative loss function, Siamese networks classifies the two input if the entered two samples are of the same class or of the other class. In [27], LeCun et al. Introduced Siamese networks as parts of their handwritten signature verification system. In a recent related study, pedestrian tracking [28], object cosegmedtation [29] showed that Siamese Networks can be used for image classification tasks and [30] captures semantics from job resumes.

3.2 the Kernel and the Range space learning

Multi Layer Perceptron (MLP) neural networks has been widely used in machine running. In general, MLP is trained by the gradient descent method and backpropagation [31]. Since the learning parameters such as learning rate or momentum value have a great impact on the performance of the gradient descent method, it is important to set these parameters carefully. However, finding the appropriate values for these parameters through the trial and error is a time-consuming task.

Recently, gradient-free learning framework for MLP has developed. By using this novel framework, the MLP is trained based on the kernel and range (KAR) space manipulation [18, 32–34]. Since this learning framework stands on linear algebra and pseudo-inverse, no parameters and no iteration are need to train the network.

Given m samples, let the training dataset is denoted by $\mathbf{X} \in \mathbb{R}^{m\times(n+1)}$ and the network output is denoted by $\tilde{Y}$. Then the MLP networks composed of $n-1$ hidden layers $\{h_1, \ldots, h_{n-1}\}$ can be represented by the following equation:

$$\tilde{Y} = \sigma\left(\left[\mathbf{1}, \sigma\left(\ldots\left[\mathbf{1}, \sigma\left(\left[\mathbf{1}, \sigma\left(\mathbf{X}\mathbf{W}_1\right)\right]\mathbf{W}_2\right)\right]\ldots\mathbf{W}_{(n-1)}\right)\right]\mathbf{W}_n\right), \tag{3.1}$$

where $\mathbf{W}_1 \in \mathbb{R}^{(n+1)\times h_1}$,$\mathbf{W}_2 \in \mathbb{R}^{(h_1+1)\times h_2}$,$\ldots, \mathbf{W}_n \in \mathbb{R}^{(h_{(n-1)}+1)\times n}$,$\mathbf{1} = [1, \ldots, 1]^T$ $\in \mathbb{R}^{m\times 1}$ and $\sigma(.)$ is activation function. We can train this network by adopting the one-hot encoded target matrix $\tilde{Y} \in \mathbb{R}^{m\times n}$ instead of network output $\tilde{Y}$, The trained weight matrices $\mathbf{W}_i$ using KAR space manipulation learning can be obtained as below [18]:

$$\begin{aligned}\mathbf{W}_i = &\left[\mathbf{1}, \sigma\left(\ldots\left[\mathbf{1}, \sigma\left(\left[\mathbf{1}, \sigma\left(\mathbf{X}\mathbf{W}_1\right)\right]\mathbf{W}_2\right)\right]\ldots\mathbf{W}_{(i-1)}\right)\right]^{\dagger}\sigma^{-1}\left(\mathbf{Y}\right),\\ &i = 1, \ldots, n.\end{aligned} \tag{3.2}$$

Chapter 4

Proposed System

4.1 Preprocessing

To be used as an input to a proposed system, the Wi-Fi signature data must be pre-processed. First, the signal has to converted to non-complex form. Due to the device firmware issues for the phase value of the 2.4Ghz Wi-Fi signal [35], we utilized only the magnitude value of the complex signal. Second, all inputs must be arranged of the same size. Since all signals have different length, we used equal length sampling to match all data to a certain length. Third, outlier issues. Wi-Fi signals are exposed to various electromagnetic noise. Among them, Burst noise [35] has a lot of unexpected effects because its value is larger than the normal signal. To minimize the effects of burst noise, we used values corresponding to the median ± 1 standard deviation were used.

4.2 Proposed Methodology

In this paper, we propose an identity verification system based on the in-air hand-written Wi-Fi signature signals (which will be called Wi-Fi signature hereafter).

In order to learn the direction-invariant deep representations for in-air signature, we used the triplet network [36] which utilizes ConvNet [37] as a feature extractor. Moreover, to achieve a faster loss convergence, we adopt the kernel and the range (KAR) space learning [18, 34] for mining the triplet input. Figure 4.1

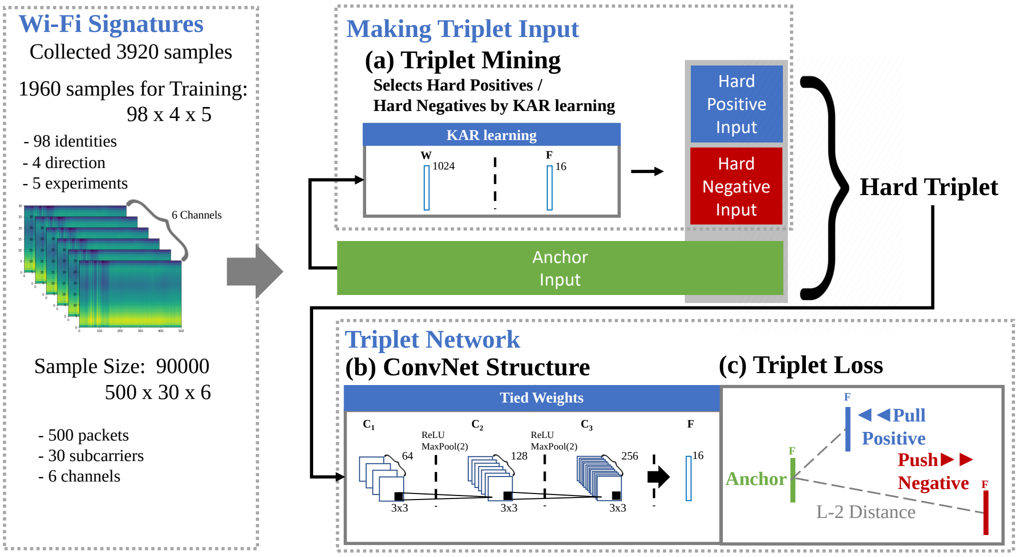

Figure 4.1: An overview of the proposed methodology.

shows an overview of the proposed system. Essentially, the KAR space projection learning is applied to mine the hard samples from the training dataset for making the triplet input (see item (a) in Fig. 4.1). The anchor sample is randomly selected from the training dataset as the reference data. The hard samples refer to which likely to be misclassified by the triplet network for a given anchor sample. Subsequently, the ConvNet structure that forms triplet network (see item (b) in Fig. 4.1) is trained based on a triplet loss function based on the distance comparison for network output vector (see item (c) in Fig. 4.1). In the following subsections, we introduce the triplet network architecture, triplet loss, and triplet mining using KAR space learning in detail.

4.2.1 Triplet loss

The purpose of the triplet loss [36] is training the ConvNet structure to learn discriminative features that allocate the samples of the same class closer and the samples of the different class far away in the feature space. The triplet input is made up of a combination of three samples, anchor sample x_0, positive sample x_+ and negative sample x_-. The anchor sample, which is the reference for the triplet input, is selected from the training data set. For selected anchor sample x_0, positive sample is selected from the same identity with that of the anchor while a negative sample is selected from different identity from that of the anchor. To make the discriminative feature vectors, we need $dist\{f(x_0), f(x_+)\}$, the distance between feature vectors of anchor $f(x_0)$ and positive sample $f(x_+)$ is larger than $dist\{f(x_0), f(x_-)\}$, distance between feature vectors of anchor and negative sample plus preset margin α. Distance measurement function is shown below:

$$dist\{f(x_0), f(x_-)\} - dist\{f(x_0), f(x_+)\} \geqq \alpha \tag{4.1}$$

By using $L2$ distance as distance function, triplet loss is calculated as below:

$$triplet\_loss = \sum_{i}^{N} max\left(\left[\|f(x_0) - f(x_+)\|_2^2 - \|f(x_0) - f(x_-)\|_2^2 + \alpha\right], 0\right), \tag{4.2}$$

Note that if $dist\{f(x_0), f(x_-)\}$ is much larger than $dist\{f(x_0), f(x_+)\} + \alpha$, the output of the loss function would be zero. In this case, it may slow down the convergence speed of the training of deep network. However, it is likely to fall into this condition if we randomly select training samples to make a triplet input. For fast loss convergence, we need to mine the triplet input that make the output of our network satisfy the condition 4.1 to ensure that the output of the loss function is non-zero.

4.2.2 Triplet Mining based on the Kernel and the Range space learning

To train the network faster, we propose training a sub-network for mining the hard positive and the hard negative sample from the training dataset. The hard positive sample is the sample which is likely to be misclassified as a negative sample by the triplet network. In other words, the distance between feature vectors of anchor and the hard positive sample is larger than other positive samples. On the other hand, the hard negative sample is likely to be misclassified as a positive sample since the distance between feature vectors of anchor and this hard negative sample is smaller than other negative samples. Hard triplet input is made by combining hard positive and hard negative samples for selected anchor samples. By using hard triplet as input to the triplet network, it is easier to satisfy the conditions of 4.1. However, we don't know which sample is the hard sample before training the triplet network. In order to make the hard triplet input before tranining the triplet network, we propose training a small sub-network before triplet network. This small sub-network is made of Multi Layer Perceptron (MLP) and we adopt the Kernel And the Range (KAR) space learning to train the MLP sub-network. As the KAR space learning has no backpropagation

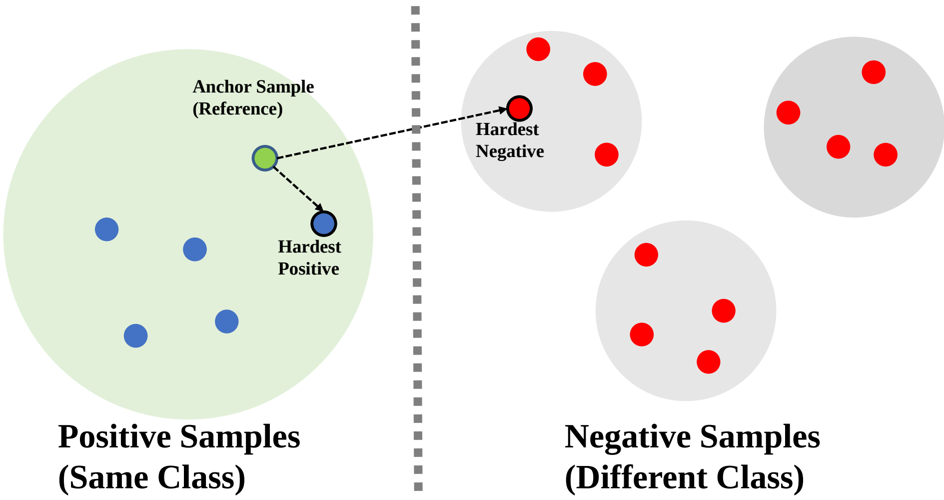

Figure 4.2: Selection of hard samples.

and no iterative learning process, we can train this sub-network with the single-shot utilizing the entire training dataset X. Given entire training dataset X, the sub-network output is given as:

$$KAR\left(\mathbf{X}\right) = \sigma\left(\left[\mathbf{1}, \sigma\left(\ldots\left[\mathbf{1}, \sigma\left(\left[\mathbf{1}, \sigma\left(\mathbf{X}\cdot\mathbf{W}_1\right)\right]\mathbf{W}_2\right)\right]\ldots\mathbf{W}_{(n-1)}\right)\right]\mathbf{W}_n\right). \quad (4.3)$$

After training the sub-network, we can mine the hard sample by measuring the $L2$ distance between every output vector of the sub-network and output vector of anchor sample. The sub-network output of a given anchor sample x_0 is $KAR(x_0)$, To mine hard-positive sample, select one sample among the sub-network output which distance to anchor feature vector $KAR(x_0)$ is larger than threshold t_+. For hard-negative sample, choose one among the sub-network output which distance to anchor feature vector is smaller than threshold t_-. Selected hard-positive and hard-negative sample satisfy this property:

$$\|KAR(\mathbf{x}_0) - KAR(\mathbf{x}_+)\|_2^2 \geq \mathrm{t}_+, \tag{4.4}$$

$$\|KAR(\mathbf{x}_0) - KAR(\mathbf{x}_-)\|_2^2 \leq \mathrm{t}_-, \tag{4.5}$$

If the hardest sample, it is known that the outlier data is likely to be selected and there is a risk of overfitting [25]. To avoid this problem, the threshold for the hard-positive and the hard-negative samples are empirically chosen as 25 and 75 percentiles of the distance between anchor and sub-network outputs.

4.2.3 The ConvNet structures

Since the three-dimensional formats of our input signature signal can be considered as an image data, we utilized deep ConvNet structure as a feature extractor. This ConvNet structure is made of three layers of ConvNet with triplet loss. Each layer of ConvNet shares their weights. Our ConvNet model is consists of three filters and output fully-connected (FC) layer as seen in (item (b) in Fig. 4.1). The depth of 3x3 sized ConvNet filters are empirically chosen as (64,128,256) with stride 1 and ReLU activation function. The size of FC layer is 16. Output FC layers are go through sigmoid activation function and normalized using $L2$ distance.

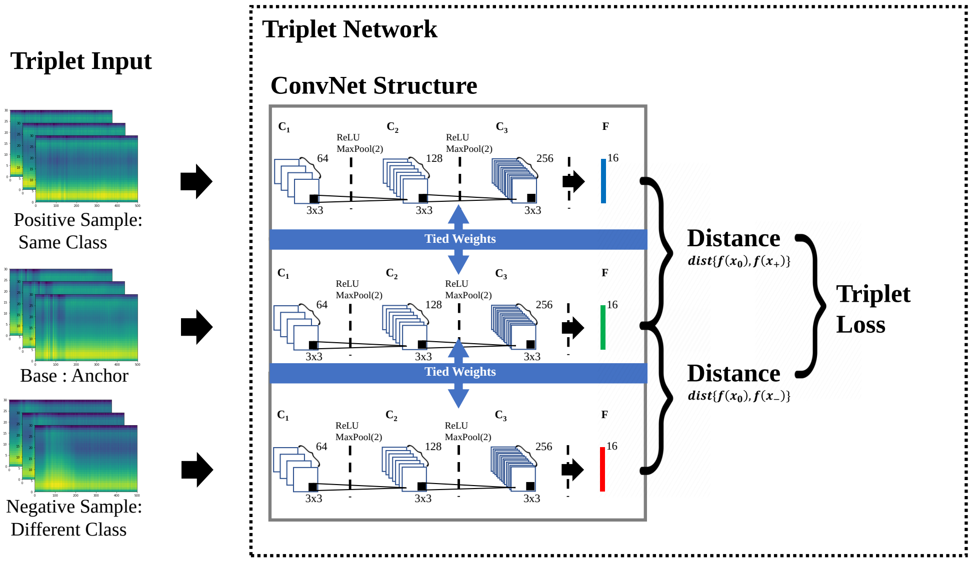

Figure 4.3: ConvNet structure.

Chapter 5

Experiments

5.1 Data set

The Wi-Fi CSI signature dataset [15] at position 1 was implemented in our experiment to measure the performance of our proposed system. These are in-house Wi-Fi signature datasets that comprise 980 signatures per direction, as shown in 5.1. The size of each sample was 500 x 30 x 6.

5.2 Experimental Settings

This section details the evaluation scenarios and experimental settings.

Table 5.1: Description of the Dataset.

Direction of signature	# of identities	# of data
1	98	980
2	98	980
3	98	980
4	98	980

5.2.1 Evaluation Scenarios

In this paper, the verification performance of the proposed method was evaluated in three ways: I) In the first experiment, verification performance was compared between the proposed system and other systems including handcraft and deep-learning based methods. II) Under the second experiments, Comparison of convergence speed was conducted between the proposed system other deep-learning based methods. III) The third experiments were conducted to compare the performance degradation between the proposed system and other deep-learning based methods when using a small-sized feature vectors. To compare the feature extraction performance of the proposed system with other deep-learning based methods, we visualized feature space into 2d euclidean space by using PCA. We compared the proposed system both with handcraft methods and deep-learning based methods. For handcraft methods, least square estimations (LSE) [38], the principal components analysis(PCA) [39] with LSE, the supprot vector machine [40] and the total error minimization with the reduced multivariate polynomal [41,42] were used. We selected parameters that perform optimally in each handcraft methods. For LSE,SVM and TER, the input signatures were reduced to 500×30 by averaging the subcarrier Axis. For PCA-LSE, input signature dimension was reduced to 40 following [15]. For SVM with Gaussian kernel function (RBF), the kernel's parameters c and γ were chosen by a grid search over the range $c \in \{0.01, 1, 10\}$ and $\gamma \in \{0.01/3000, 0.1/3000, 1/3000, 10/3000, 100/3000\}$. For TER, parameter M is chosen among $M \in \{1, 2, 3\}$ and set $\tau = \eta = 0.5$ following [42]. For com-

Table 5.2: The network structure of KAR space learning.

Layer	Size	Activation
Input	500×30×6	
Fully-Connected 1	1×1×1024	$\sigma = tan^{-1}$
Fully-Connected 2	1×1×16	$\sigma = tan^{-1}$
Output	1×1×50	

parison with deep-learning based methods, Siamese network [26] and baseline triplet network [36] were used. The verification performance was evaluated in terms of the Equal Error Rate (EER, %). We implemented a random 5-runs of 2-fold cross-validation tests. Due to the hardware memory limitations, we used downsized negative pairs according to the number of positive data pairs for calculating the EER. The size of positive data pairs and downsized negative pairs were 18620.

5.2.2 Parameter Settings

For the proposed system, the structure of the KAR learning MLP sub-networks is shown in 5.2. We empirically chose two layers size of 1024 and 16. The size of the second layer was equivalent as feature vectors of the proposed ConvNet. The weights in the layers were initialized as a normal distribution of 0 to 1 before training. We used tan^{-1} as an active function following [32]. We used the same ConvNet structure shown in 5.3 and parameter settings for the proposed system and the deep-learning based methods. The CovNet structure consists of 3 convolution filters size of 3×3 and stride 1. A ReLU activation function and 2×2 max-pooling layers is applied between the filters. The depth of each

Table 5.3: The structure of ConvNet model.

Layer	Activation	Kernel / Stride	Input Size
Conv 1	ReLU	(3×3)×64/1	500×30×6
MaxPool 1		(2×2)/1	500×30×64
Conv 2	ReLU	(3×3)×128/1	250×15×64
MaxPool 2		(2×2)/1	250×15×128
Conv 3	ReLU	(3×3)×256/1	125×8×128
MaxPool 3		(2×2)/1	125×8×256
Fully-Connected	Sigmoid	16	63×4×256
L-2 Norm			1×1×16
Concat			1×1×16

layer was chosen as {64,128,256}. The output layer with sigmoid activation was regularized using $L2$ penalty of 0.0001. The size of the final feature vectors was 16. The parameter settings for traning the deep learing network was learning rate of 0.0005, the number of iteration as 3000, and set the mini-batch size as 32. We used adam optimatizer for calculating the loss function. We initialized the ConvNet structures following [26] before training. For convolution filters, we used a normal distribution of 0 mean and standard deviation of 0.0001. For the biases, parameters for normal distribution was 0.5 mean and standard deviation of 0.01. For calculating triplet loss, we set the alpha value as 0.5.

5.3 Experimental Results

5.3.1 Performance

For the first experiment, 5.4 shows the average EER from 5-runs of 2-fold cross-validation tests under the optimal parameter settings. As shown in Table 5.4, the proposed system shows the best performance among the handcraft and deep-learning based methods with 19.35% EER. Deep learning based methods performed better performance than handcraft methods since they were able to utilize the entire input signal, and their ability to feature extraction were better than handcraft methods.

Table 5.4: Performance benchmarking with respect to the best EER (%) averaged from five runs of two-fold cross-validation test on Wi-Fi CSI signature dataset.

Methology	Best EER (%)	Condition
LSE	48.44	-
PCA-LSE	30.79	Reduced dimension=40
SVM (Linear)	28.23	c=1
SVM (RBF)	24.31	c=1, γ=0.01/3000
TER-RM2	35.84	M=1,τ=η=0.5
Siamese network	23.53	lr=0.00005
Baseline triplet network	20.34	lr=0.00005, α=0.1
Proposed system	**19.35**	**lr=0.00005, α=0.1**

As shown in 5.1, the proposed system also showed the widest Area under ROC curve (AUC).

5.3.2 Convergence speed

5.2 shows the convergence of the normalized loss function during the training. The convergence speed of the proposed system was the fastest among the deep learning-based methods, meaning that KAR running has accelerated the learning speed.

5.3.3 Effect of the size of Feature Vector

Figure 5.3 shows EER for different sizes of feature vectors in the feature space. In general, in metric learning systems, the smaller the size of the feature vectors,

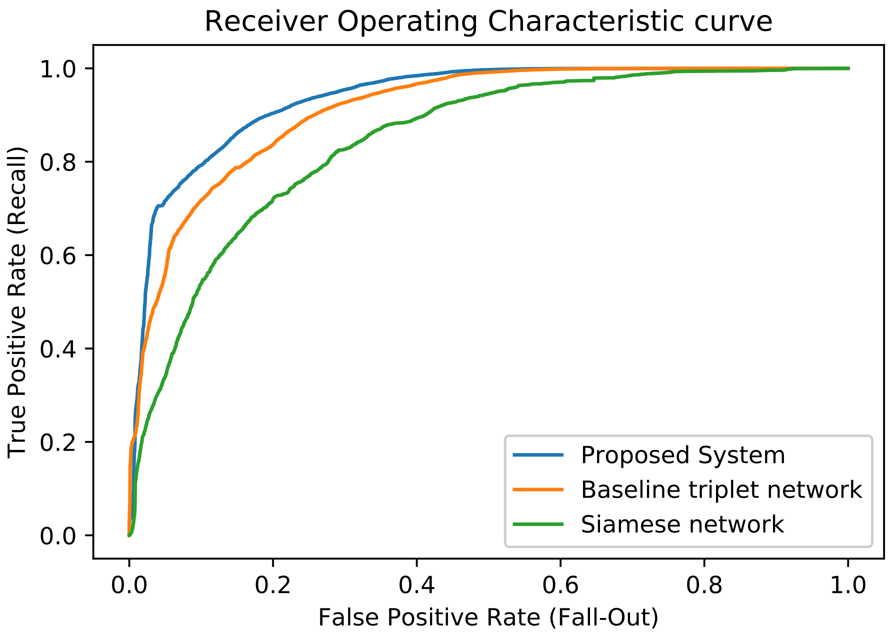

Figure 5.1: Normalized training loss curve.

Size of the Feature Vector	16	8	4	2
Siamese network	20.34	22.05	29.39	34.26
Baseline triplet network	18.37	19.52	19.20	24.76
Proposed system	16.92	18.03	18.08	26.64

the less performance of the system. From the results, we can see the proposed system shows less performance degradation by a small size feature vector.

5.3.4 2D visualization of the feature vector

To visualize the performance of feature extraction property of the proposed system, we project the feature space to two-dimensional euclidean space by using PCA in Figs. 5.6,5.4,5.5. Through the feature space projected in the 2d euclidean space, we were able to intuitively verify the feature extraction performance of

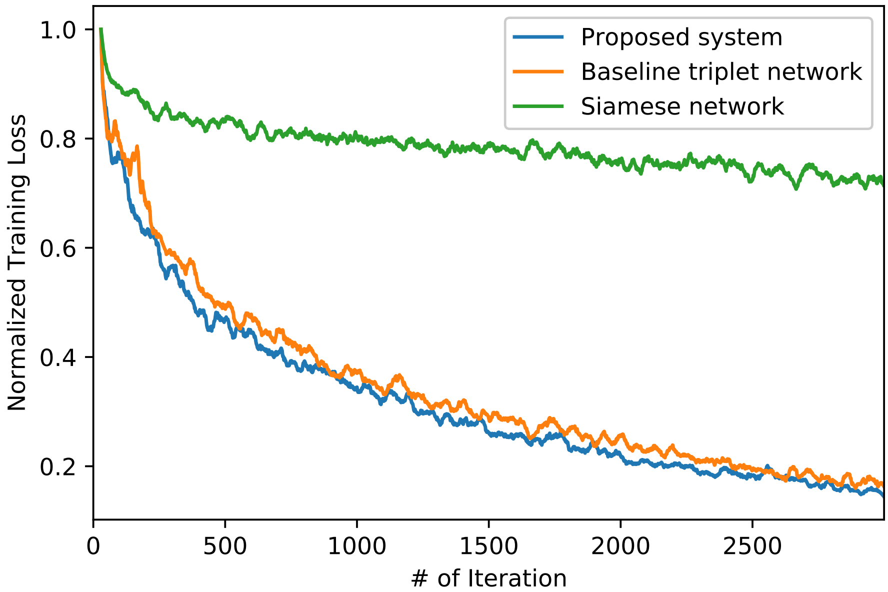

Figure 5.2: normalized training loss trends

each system by making sure that the point of being close to the space belongs to the same identity. To easily check the distance between points in 2d space, we marked only the first 10 identities. As shown in the figure, the proposed system used the 2d euclidean space most efficiently. The Siamese network linearly arranged feature vectors, making it difficult to distinguish between classes, and in the case of the baseline triplet network, there were many overlapping areas between identities.

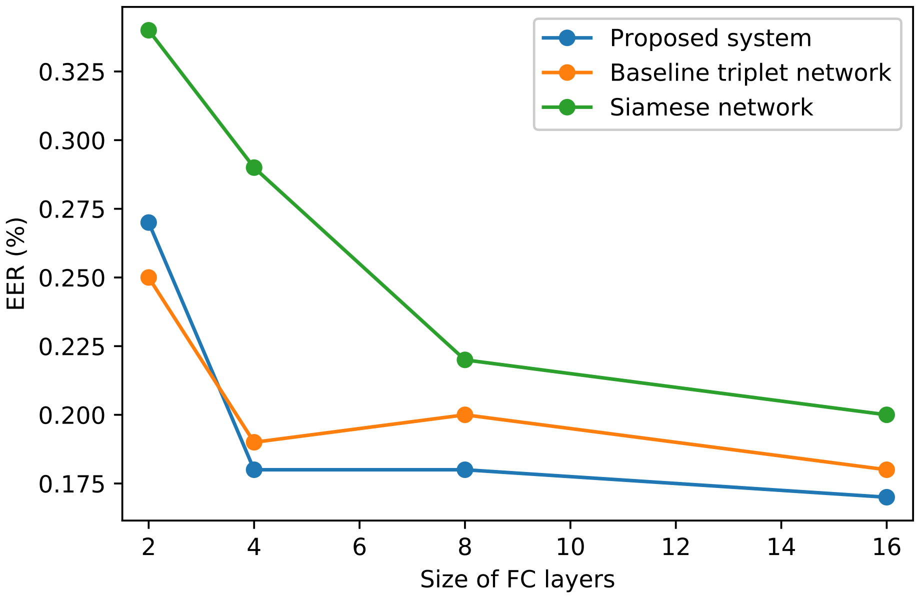

Figure 5.3: Size of FC layer.

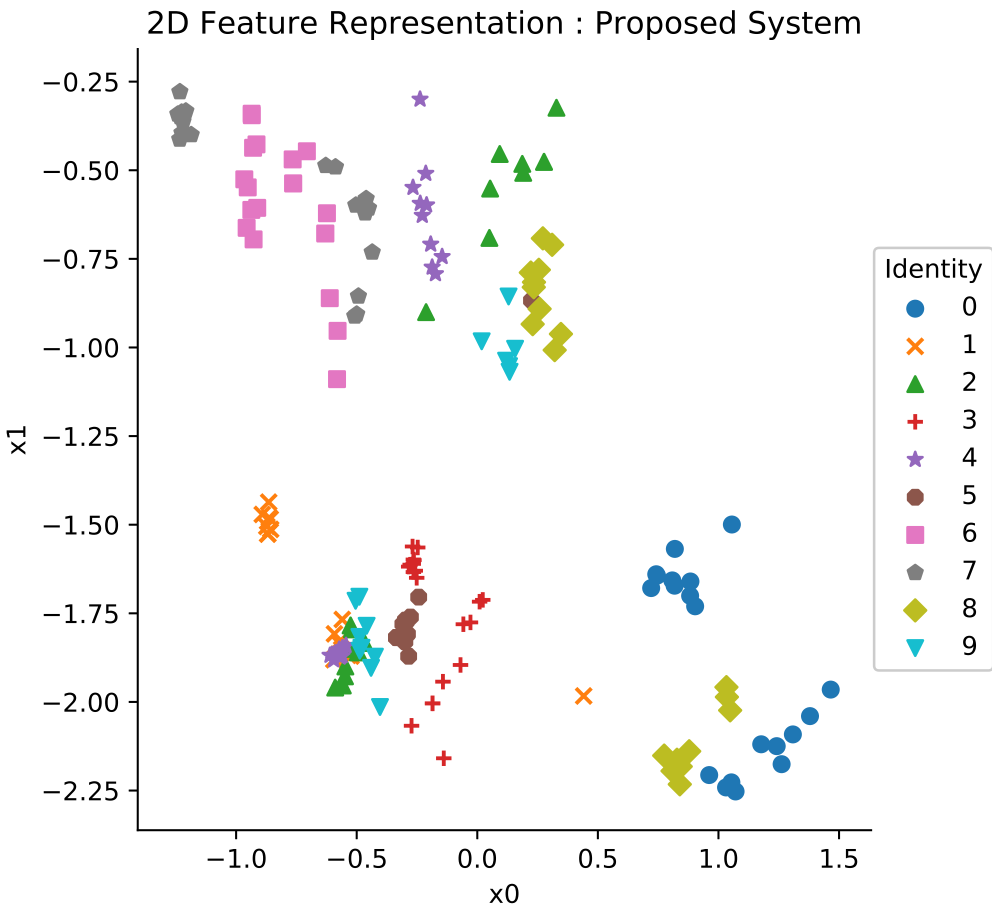

Figure 5.4: 2D Feature Representation : Proposed System.

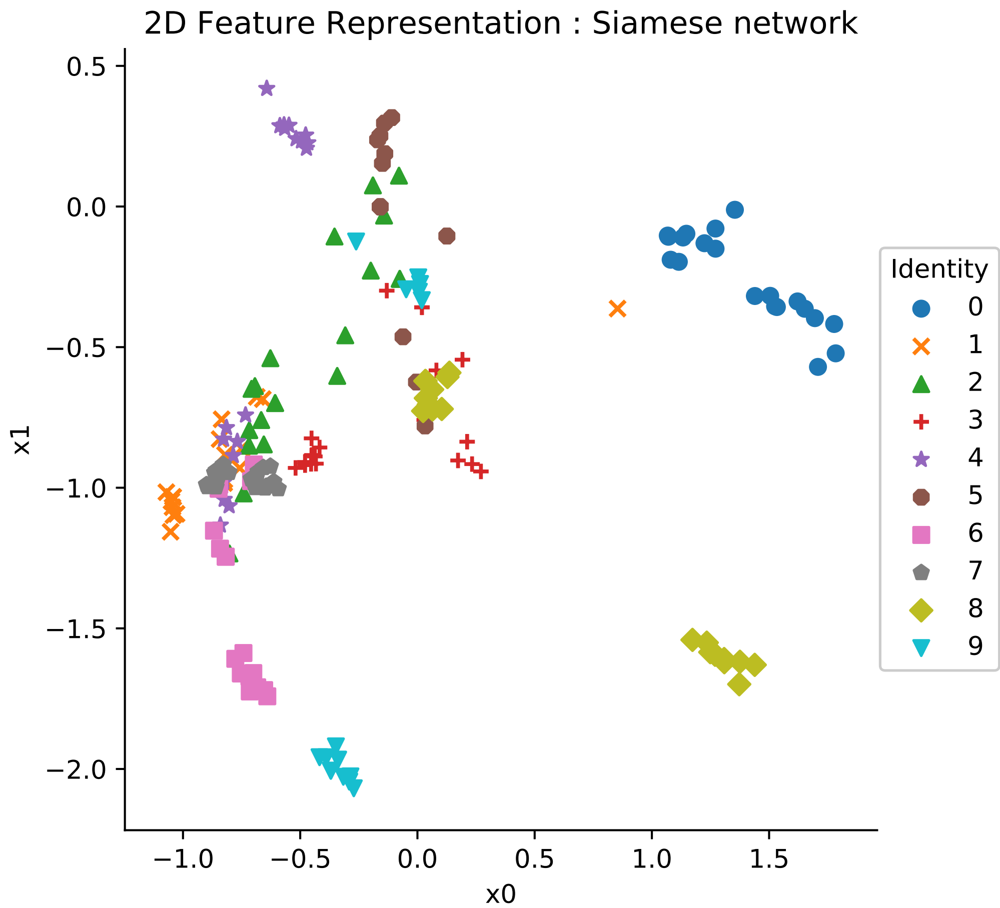

Figure 5.5: 2D Feature Representation : Baseline triplet network.

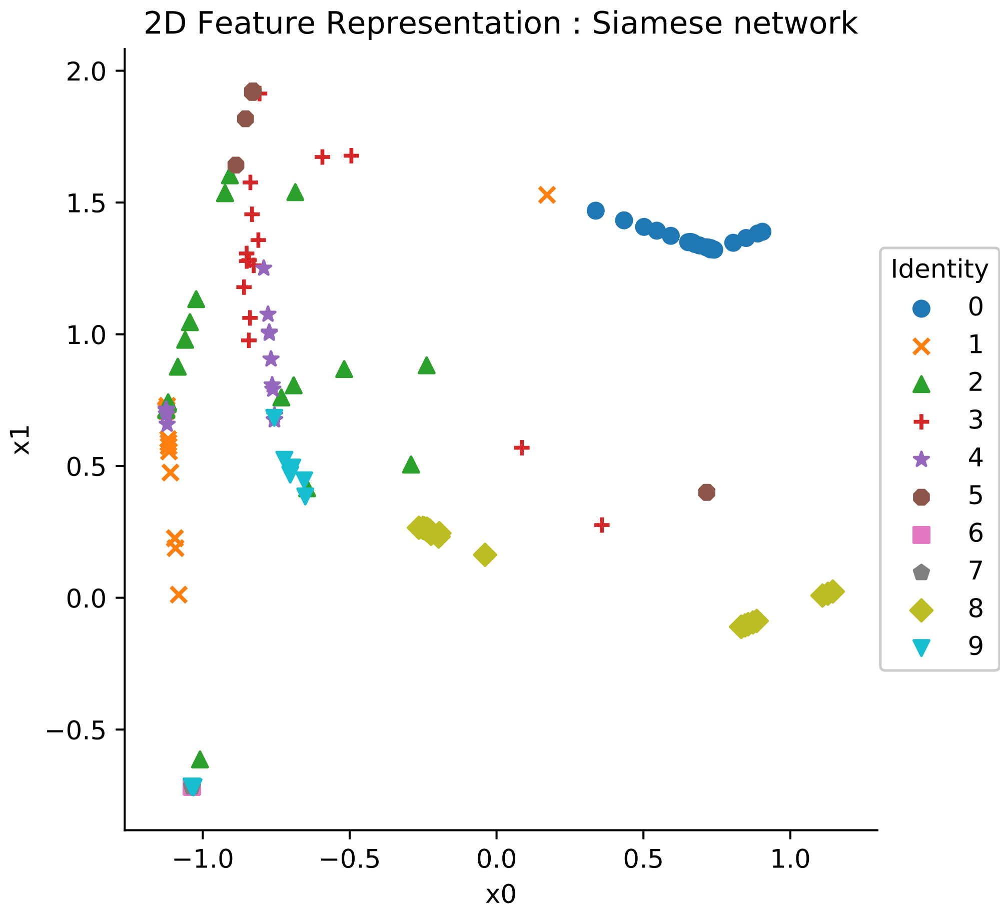

Figure 5.6: 2D Feature Representation : Siamese network.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

In this paper, we proposed a system for identity verification from Wi-Fi in-air signature signals based on triplet network. Three-channel ConvNet structures was adopted in order to learn discriminative features from the in-air signatures. Moreover, we proposed a input triplet mining approach based on the kernel and range space learning to faster the convergence speed. Our experimental results on in-house Wi-Fi handwritten signature dataset showed the proposed network outperforms handcraft methods and Siamese network by improved verification accuracy and faster loss convergence.

Bibliography

[1] M. M. Fahmy, "Online handwritten signature verification system based on dwt features extraction and neural network classification," *Ain Shams Engineering Journal*, vol. 1, no. 1, pp. 59–70, 2010.

[2] J. Galbally, M. Diaz-Cabrera, M. A. Ferrer, M. Gomez-Barrero, A. Morales, and J. Fierrez, "On-line signature recognition through the combination of real dynamic data and synthetically generated static data," *Pattern Recognition*, vol. 48, no. 9, pp. 2921–2934, 2015.

[3] A. Sanmorino and S. Yazid, "A survey for handwritten signature verification," in *the 2nd International Conference on Uncertainty Reasoning and Knowledge Engineering.* IEEE, 2012, pp. 54–57.

[4] E. Sesa-Nogueras, M. Faundez-Zanuy, and J. Mekyska, "An information analysis of in-air and on-surface trajectories in online handwriting," *Cognitive Computation*, vol. 4, no. 2, pp. 195–205, 2012.

[5] D. Hutton, "Biometrics: Identity verification in a networked world," *Kybernetes*, 2004.

[6] G. Bailador, C. Sanchez-Avila, J. Guerra-Casanova, and A. de Santos Sierra, "Analysis of pattern recognition techniques for in-air signature biometrics," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2468–2478, 2011.

[7] J.-H. Jeon, B.-S. Oh, and K.-A. Toh, "A system for hand gesture based signature recognition," in *the 12th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, 2012, pp. 171–175.

[8] J. Malik, A. Elhayek, S. Ahmed, F. Shafait, M. Malik, and D. Stricker, "3DAirSig: A framework for enabling in-air signatures using a multi-modal depth sensor," *Sensors*, vol. 18, no. 11, p. 3872, 2018.

[9] H. Ketabdar, P. Moghadam, B. Naderi, and M. Roshandel, "Magnetic signatures in air for mobile devices," in *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services companion*. ACM, 2012, pp. 185–188.

[10] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "Wfid: Passive device-free human identification using WiFi signal," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. ACM, 2016, pp. 47–56.

[11] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 267–276.

[12] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using wifi channel state information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017.

[13] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous wifi-based gesture recognition system," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2015, pp. 1472–1480.

[14] R. Nandakumar, B. Kellogg, and S. Gollakota, "Wi-fi gesture recognition on existing devices," *arXiv preprint arXiv:1411.5394*, 2014.

[15] H.-C. Moon, S.-I. Jang, K. Oh, and K.-A. Toh, "An in-air signature verification system using Wi-Fi signals," in *Proceedings of the 4th International Conference on Biomedical and Bioinformatics Engineering*. ACM, 2017, pp. 133–138.

[16] A. Pokkunuru, K. Jakkala, A. Bhuyan, P. Wang, and Z. Sun, "Neuralwave: Gait-based user identification through commodity WiFi and deep learning," in *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2018, pp. 758–765.

[17] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT," in *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2017, p. 5.

[18] K.-A. Toh, Z. Lin, Z. Li, B. Oh, and L. Sun, "Gradient-free learning based on the kernel and the range space," *arXiv preprint arXiv:1810.11581*, 2018.

[19] L. Yang and R. Jin, "Distance metric learning: A comprehensive survey," *Michigan State Universiy*, vol. 2, no. 2, p. 4, 2006.

[20] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Deep metric learning for person re-identification," in *2014 22nd International Conference on Pattern Recognition*. IEEE, 2014, pp. 34–39.

[21] K. Q. Weinberger, J. Blitzer, and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," in *Advances in neural information processing systems*, 2006, pp. 1473–1480.

[22] D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng, "Person re-identification by multi-channel parts-based cnn with improved triplet loss function," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1335–1344.

[23] S. Ding, L. Lin, G. Wang, and H. Chao, "Deep feature learning with relative distance comparison for person re-identification," *Pattern Recognition*, vol. 48, no. 10, pp. 2993–3003, 2015.

[24] F. Wang, W. Zuo, L. Lin, D. Zhang, and L. Zhang, "Joint learning of single-image and cross-image representations for person re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1288–1296.

[25] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815–823.

[26] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *ICML deep learning workshop*, vol. 2, 2015.

[27] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a" siamese" time delay neural network," in *Advances in neural information processing systems*, 1994, pp. 737–744.

[28] L. Leal-Taixe, C. Canton-Ferrer, and K. Schindler, "Learning by tracking: Siamese cnn for robust target association," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2016.

[29] P. Mukherjee, B. Lall, and S. Lattupally, "Object cosegmentation using deep siamese network," *arXiv preprint arXiv:1803.02555*, 2018.

[30] S. Maheshwary and H. Misra, "Matching resumes to jobs via deep siamese network," in *Companion of the The Web Conference 2018 on The Web Conference 2018*. International World Wide Web Conferences Steering Committee, 2018, pp. 87–88.

[31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[32] K. Toh, "Analytic network learning," *arXiv preprint arXiv:1811.08227*, November, 2018.

[33] K.-A. Toh, "Kernel and range approach to analytic network learning," *International Journal of Networked and Distributed Computing*, vol. 7, no. 1, pp. 20–28, December 2018.

[34] ——, "Learning from the kernel and the range space," in *the IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS).* IEEE, 2018, pp. 1–6.

[35] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking.* ACM, 2015, pp. 65–76.

[36] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *International Workshop on Similarity-Based Pattern Recognition.* Springer, 2015, pp. 84–92.

[37] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[38] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification.* John Wiley & Sons, 2012.

[39] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[40] V. Vapnik, *The nature of statistical learning theory.* Springer science & business media, 2013.

[41] K.-A. Toh, "Fingerprint and speaker verification decisions fusion," in *12th International Conference on Image Analysis and Processing, 2003. Proceedings.* IEEE, 2003, pp. 626–631.

[42] K.-A. Toh and H.-L. Eng, "Between classification-error approximation and weighted least-squares learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 658–669, 2008.

국문요약

와이파이 신호를 이용한 공중 서명 인식 시스템

…

핵심되는 말: …