Improving the Triplet Network for Wi-Fi Based Handwritten Signature Verification

Young-Woong Kwon

The Graduate School
Yonsei University
School of Electrical and Electronic Engineering

Improving the Triplet Network for Wi-Fi Based Handwritten Signature Verification

A Masters Thesis
Submitted to the School of Electrical and Electronics 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

First of all, I would like to thank Professor Kar-Ann Toh for steering the direction of my research. His warm encouragement and careful guidance were very helpful for me. Thank you for my director Kiheon Kim and my manager Jongman Kim for recommending me to study at graduate school. Thanks for my company, Samsung SDI for supporting me to study at graduate school. I also want to give my special thanks to Jooyoung, Kim. His consideration and positive attitude always motivated me to do better. Thank you to all my lab members for helping to enjoy school life. Last but not least, I couldn't have finished this work without my dear wife, who always trusted and waited for me despite going back to graduate school at a late age.

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 challenging task in signature verification since the shape of the signal varies according to the direction in which the signature is written. By using the traditional methods, it was difficult to verify identity from signatures entered in various directions. Moreover, limited size of the dataset also limited the training of deep learning models. In this paper, we propose a method for identity verification from Wi-Fi in-air signatures based on triplet network. Three-channel ConvNet structures is adopted to learn discriminative features from relatively small size of in-air signature datasets. 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 the Wi-Fi CSI signature dataset shows that the proposed method out-outperforms both the traditional and deep learning based methods by improved verification accuracy and faster loss convergence.

Key words : 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

Written signatures have long been used as an identification tool and various methods for verifying individuals base on them have been developed [1–4]. The use of biometric identity verification is growing with the popularity of mobile devices. Biometric identity verification systems are more secure, convinient, and provide greater accountability than traditional complex passwords [5]. In-air signatures have become possible with the development of specialized sensors, such as depth cameras and magnetic sensors [6–9]. In-air signatures can be made using Wi-fi Channel State Information (”CSI”) signals. Typically people make a certain gesture when signing their signature on a piece of paper. With in-air signatures, people simply make the same gesture in the air. Their movements are detected by Wi-fi receivers with appropriate software installed. The software detects changes in the Wi-fi signals caused by hand movements. This technique uses widespread devices, so it does not require the use of any special input devices. [10, 11] used CSI signals to recognize users’ gestures and [12] used them to specifically identify signatures written in the air. However, due to limitations of the traditional feature extractor, [12] was only able to identify in-air signatures entering from the specific direction. This makes the user difficult to input their signatures. Recent studies have been conducted on the feasibility of using deep learning technology to characterize CSI signals. Deep learning-based models are popular often studied as solutions to this situaiton because of their automated feature extractors and the fact that they have superior classification capabilities than traditional models. Deep learning has been used to recognize users based on their body shape [13] and behavior [14]. In this thesis, deep learning technology was used to create a Wi-Fi CSI system that can recognize a user's identity with multi-direction in-air signatures. The deep learning model tested in this study used a triplet network to increase classification and feature extraction accuracy, and to improve the model's convergence speed using the kernel and range space learning techniques [15].

1.2 Motivation and Contributions

The main contributions of this thesis were

- Proposing a system to verify user identities using Wi-Fi handwritten signature signals using a deep triplet network;
- Using the kernel and range (KAR) space learning to mine distinctive triplet inputs which boost convergence speed and reduce triplet network training loss; and
- Empirically testing the proposed system on a dataset of Wi-Fi handwritten signatures collected from 98 subjects.

1.3 Paper Organization

The rest of this paper is organized as follows. Section 2 discusses related works about triplet networks and KAR space learning. Section 3 discusses the proposed system. Section 4 describes this thesis's experimental and analysis results. Section 5 concludes the thesis.

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 converts multi-dimensional data to group data in a feature space. It has been used for computer vision tasks such as image classification and content-based image retrieval [16]. In the past, features were extracted and similarity was measured using traditional algorithms like HOG and LBP. Rescently, deep learning-based methods have been widely used instead as they have enabled feature extraction and metric learning in a single framework [17]. The deep networks automatically learns image classification features. Deep learning-based models include the Siamise network and the Triplet network.

3.1.1 Triplet Network

Triplet network is a metric learning model which group triplet data in a feature space [16, 18]. It is widely used to identify a person based on several images in a process known as person re-identification, which is similar to butis more challanging than biometric identification due to the low quality and high variety of input images [17]. Other factors, such as changing clothes, poses, and image angles makes it difficult to identify the target person in question.

Triplet data is composed of anchor, positive, and negative data points. Training of the triplet network involves the creation of feature vectors to be placed in the appropriate feature space in which positive data is close to the anchor and negative data is far from the anchor. It may not be necessary to train an algorithm using a large number of triplets. Training can be optimized by using only the most learning-efficient triplet. [19–21] used triplet networks with used triplets for only a small number of classes, which were used in a random order. Recently, [22] used triplet mining to speed up convergence. They selected inputs from a large set at each training iteration using the network. However, these sets required the availability of a large amount of training data. This type of training requires the use of large amounts of data separated into only a few classes. Empirical data is not suitable for this type of training strategy because there is usually not enough of it and it is divided into too many classes. The kernel and range space manipulation methods were used for training to reduce the number of required triplets.

3.1.2 Siamese Networks

Siamese neural networks consist of twin networks which accept distinct inputs but are joined by an energy function at the top [23]. By using a constrative loss function, Siamese networks determine whether two inputs are in the same class. LeCun et al. introduced Siamese networks as parts of their handwritten signature verification system. Recent studies used Siamese networks to track pedestrians [24], group objects [25], and [26] capture information from resumes, indicating that they are suitable for image classification tasks.

3.2 the Kernel and the Range space learning

Multi Layer Perceptron (MLP) neural networks have been widely used in machine learning. In general, MLPs are trained using the gradient descent and backpropagation methods [27]. Learning parameters, such as learning rate and momentum value, have a significant impact on gradient descent performance, they must be set carefully. However, finding the best values through trial and error is time-consuming.

Recently, gradient-free learning frameworks for MLP have been developed which rely on KAR space manipulation [15, 28–30]. This learning framework is based on linear algebra and pseudo-inverse functions, so it does not require any parameters or iterations.

Given m samples, the training dataset can be denoted by $\mathbf{X} \in \mathbb{R}^{m\times(n+1)}$ and the network output can be denoted by $\tilde{Y}$. An MLP network 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. This network can be trained by adopting the one-hot encoded target matrix $\tilde{Y} \in \mathbb{R}^{m\times n}$ instead of network output $\tilde{Y}$. The weighted matrices $\mathbf{W}_i$ trained using KAR space manipulation learning can be written as follows [15]:

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

Chapter 4

The proposed System

4.1 Preprocessing

Wi-Fi signature data must be pre-processed to be used as an input in the proposed system. First, data must be simplified. In this study, only the magnitude of the complex signal was utilized due to device firmware issues with the phase value of the 2.4Ghz Wi-Fi signal [31]. Second, the data must be made to be the same size. All signals have different lengths, so equal length sampling was used to ensure that all data had the same length. Third, outliers must be addressed. Wi-Fi signals are interrupted by other electromagnetic radiation of which burst noise [31] can have significant effects. Thus, only data within one standard deviation of the median were analyzed to minimize the effects of burst noise.

4.2 Proposed Methodology

This thesis proposed an identity verification system based on the Wi-Fi signals produced by in-air handwritten signatures (”Wi-Fi signature”). Triplet networks which utilized ConvNet as the feature extractor were used to learn the directioninvariant deep representations of in-air signatures [32]. KAR space learning was used to mine triplet inputs to achieve faster loss convergence [15, 30]. Figure 4.1

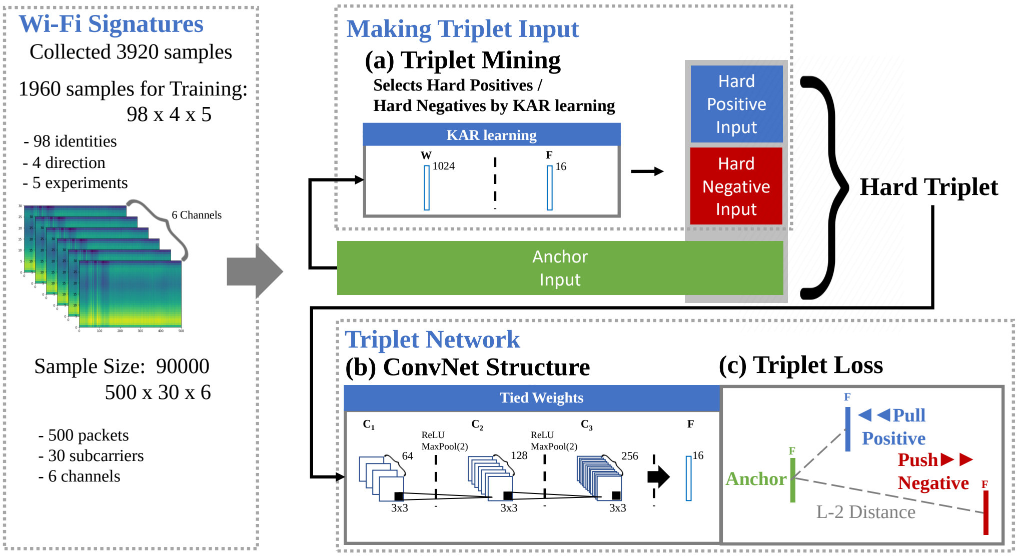

Figure 4.1: An overview of the proposed methodology.

shows an overview of the proposed system. KAR space projection learning is used to mine hard samples from the training dataset to create triplets (Fig. 4.1(a)). The anchor sample is randomly selected from the training dataset. The hard samples are those which are likely to be misclassified by the triplet network for a given anchor sample. After the data is selected, the triplet network (Fig. 4.1(b)) is trained based on the network output vector distance comparison (Fig. 4.1(c)). The following subsections discuss the triplet network architecture, triplet loss, and triplet mining using KAR space learning.

4.2.1 Triplet Loss

Triplet loss [32] trains the ConvNet structure to learn the features that define data position in the feature space. Triplet inputs are composed of a combination of three samples, an anchor sample x_0, a positive sample x_+ and a negative sample x_-. The anchor sample, the reference for the triplet input, is selected from the training data set. The positive sample has the same identity as the anchor sample while the negative sample has a different identity. $dist\{f(x_0), f(x_+)\}$, the distance between the feature vectors of anchor $f(x_0)$ and positive sample $f(x_+)$ is larger than $dist\{f(x_0), f(x_-)\}$, the distance between feature vectors of the anchor and the negative sample plus a preset margin α to produce discriminative feature vectors. The distance measurement function is:

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

By using the $L2$ distance as the distance function, triplet loss is defined as:

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

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 is zero, significantly slowing deep network convergence. This condition is likely to occur if triplets are composed by randomly selecting training samples. Thus, triplet inputs must be selected that cause the loss function to produce a non-zero result.

4.2.2 Triplet Mining Based on KAR Space Learning

To train the triplet network faster, a sub-network for mining the hard positive and negative samples from the training dataset was trained. The hard positive sample is likely to be misclassified as a negative sample by the triplet network (Fig. reffig2). The distance between the feature vectors of the anchor and hard positive samples is larger than other positive samples. The hard negative sample is likely to be misclassified as a positive sample because the distance between the feature vectors of the anchor and the hard negative sample is smaller than the difference between the feature vectors of the anchor and other negative samples. Hard triplet inputs are made by combining hard positive and hard negative samples with selected anchor samples. Using hard triplets as triplet network inputs more easily, satisfies 4.1. However, before training the triplet network, it is impossible to identify hard samples. In order to make hard triplets before tranining the triplet network, a smaller sub-network was trained before training the main triplet network. This smaller sub-network was made of an MLP and was trained using KAR space learning. As KAR space learning has no backpropagation and no iterative learning process, it was trained all at once on the entire training dataset X. The sub-network output was defined 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, the hard samples were mined by measuring the $L2$ distance between every output vector of the sub-network and the output vector of anchor sample. The sub-network output for a given anchor sample x_0 was

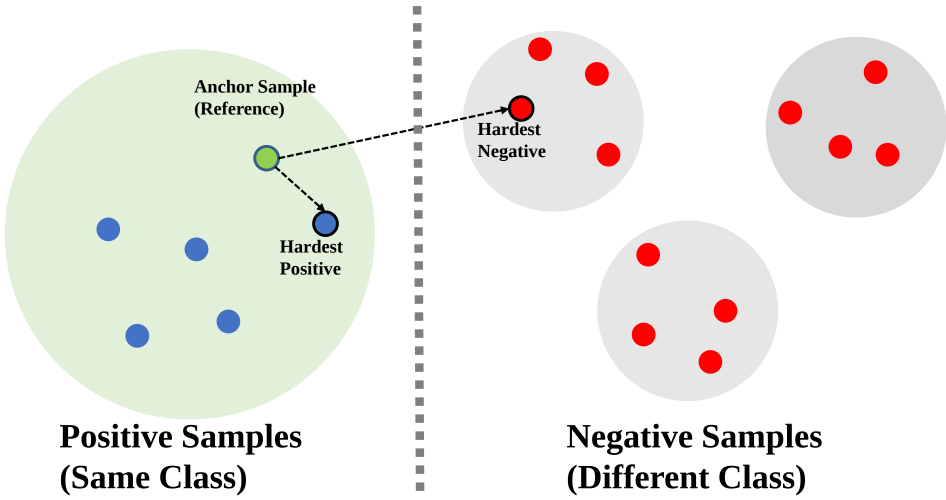

Figure 4.2: Selection of hard samples.

$KAR(x_0)$. To mine hard-positive samples, one sample was selected from among the sub-network outputs for which the distance to the anchor feature vector $KAR(x_0)$ was larger than t_+. Hard-negative samples were chosen, from among the sub-network outputs for which the anchor feature vector was smaller than t_-. The selected hard-positive and hard-negative samples satisfied the following properties, respectively:

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

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

If the hardest sample, outlier data is more likely to be selected than other data, resulting in an increased risk of overfitting [22]. To avoid this problem, the threshold for the hard-positive and the hard-negative samples were empirically chosen as the 25th and 75th percentiles of the distance between the anchor and sub-network outputs.

4.2.3 ConvNet Structures

The three-dimensional format of the input signature signal is similar to that of image data, so deep ConvNet structures were used as the feature extractor. The ConvNets used in this thesis were made of three layers of ConvNets with triplet loss. Each layer of ConvNet shared their weights. The ConvNets in this

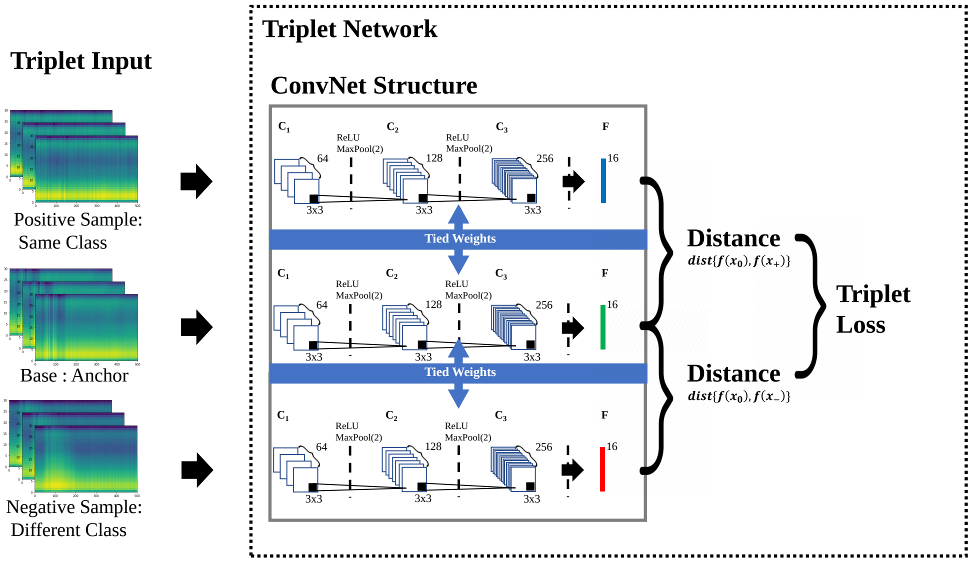

Figure 4.3: ConvNet structure.

study consisted of three filters and a fully-connected output layer (Fig. 4.3). The depth of the ConvNet filters was set at 64,128,256 with stride 1 and ReLU activation functions. The size of the fully-connected layer was 16. The fully-connected output layers were processed through sigmoid activation functions and were normalized according to the $L2$ distance.

Chapter 5

Experiments

5.1 Data set

The Wi-Fi CSI signature dataset [12] at position 1 was used to test the proposed system. These dataset contained 10 signatures made by 98 people in each direction for a total of 980 signatures made in each direction. The size of each sample had a data resolution of 500 x 30 x 6.

5.2 Experimental Parameters

The proposed method's performance, convergence speed, and degradation were compared to those of other traditional methods. To compare the proposed method's feature extraction performance to that of traditional and deep-learning based methods, feature space was visualized as a 2D Euclidean plane using principal component analysis ("PCA"). The proposed method's and traditional methods and deep-learning based methods. For handcraft methods' least square estimations ("LSE") [33], PCA with LSE [34], support vector machine [35] and the total error minimization with reduced multivariate polynomials [36, 37] were compared. Parameters were selected that performed optimally in each traditional method. For LSE,SVM, and TER, the input signatures were reduced to 500×30 by averaging the subcarrier axis. For PCA-LSE, the input signature dimension was reduced to 40 following [12]. For SVM with a 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, the parameter M was chosen from among $M \in \{1, 2, 3\}$ and $\tau = \eta = 0.5$ following [37]. For comparison with deep-learning based methods, Siamese network [23] and baseline triplet network [32] were used. The proposed method was compared to Siamese networks [23] and baseline triplet networks [32] as deep learning based methods. Verification performance was evaluated in terms of the equal error rate ("EER"). Random five-runs of two-fold cross-validation tests were conducted. Due to hardware memory limitations, the number of negative pairs used was reduced to the number of positive data pairs used for calculating the EER for a total of 18,620 pairs. The structure of the KAR learning MLP sub-networks in the proposed system is shown in 5.1. The two layers' sizes were set to 1,024 and 16. The size of the second layer was equal to the number of feature vectors of the proposed ConvNet. The weights in the layers were initialized in a normal distribution between 0 and 1 before training. tan^{-1} was used as an active function following [28]. The same ConvNet structure shown in 5.2 and parameters were used for the proposed and deep learning based methods. The CovNet structure consisted of three 3×3 convolution filters and stride one. A ReLU activation function and 2×2 max-pooling layers were applied between the

Table 5.1: KAR space learning network structure.

Layer	Size	Activation Function
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	

Table 5.2: ConvNet model structuer.

Layer	Activation Function	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

filters. The depth of each layer was set to {64,128,256}. The output layer with sigmoid activation was regularized according to $L2$ with a penalty of 0.0001. The size of the final feature vectors was 16. The deep learing network was trained with a learning rate of 0.0005, 3,000 iteration, and a mini-batch size of 32. The Adam optimizer was used to calculate the loss function. The ConvNet structures were initialized before training following [23]. The convolution filters used a normal distribution with a mean of 0 and a standard deviation of 0.0001. The biases used normal distribution with a mean of 0.5 and a standard deviation of 0.01. Triplet loss was calculated with an alpha value of 0.5.

5.3 Experimental Results

5.3.1 Performance

5.3 shows the average EER from the first experiment generated by five-runs of two-fold cross-validation tests under optimal parameter settings. The proposed method had an EER of 19.35%, out-performing both the traditional and deep learning based methods (Table 5.3). Deep learning based methods out-performed the traditional methods because they utilized the entire input signal and had better feature extraction abilities.

Table 5.3: Average EER of five-runs of two-fold cross-validation tests.

Methodology	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**

The proposed method had the widest area under the receiver operating characteristic ("ROC") curve (Fig. 5.1).

5.3.2 Convergence Speed

Fig. 5.2 shows the convergence of the normalized loss function during training. The proposed method had a faster convergence speed than the deep learning based methods, indicating that KAR learning accelerated the learning speed.

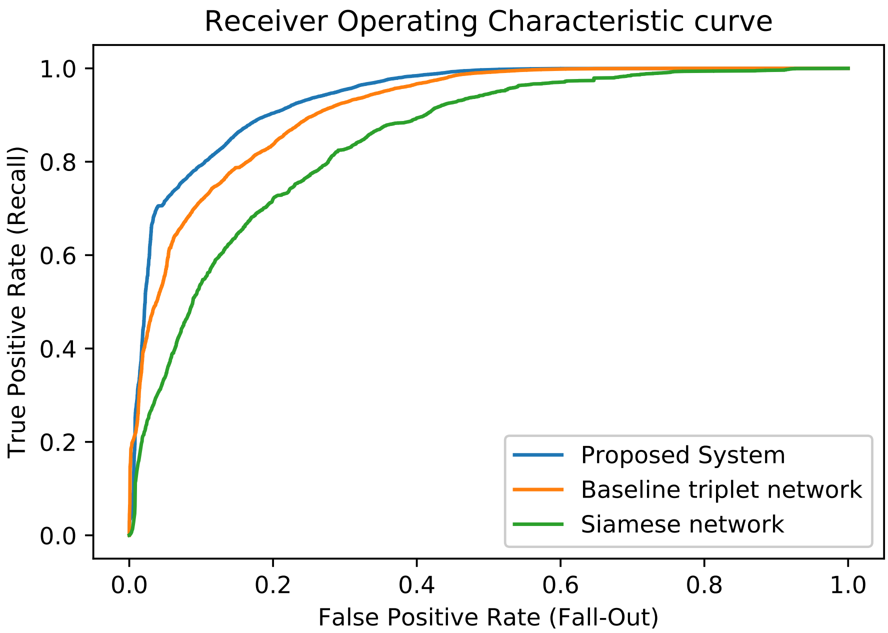

Figure 5.1: Normalized training loss curve.

Feature vector size	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 method	16.92	18.03	18.08	26.64

5.3.3 Feature Vector Size Effect

Fig. 5.3 shows the EERs for different sizes of feature vectors in the feature space. For metric learning systems, feature vector size is generally positively correlated with system performance. The proposed system's performance was negatively correlated with feature vector size.

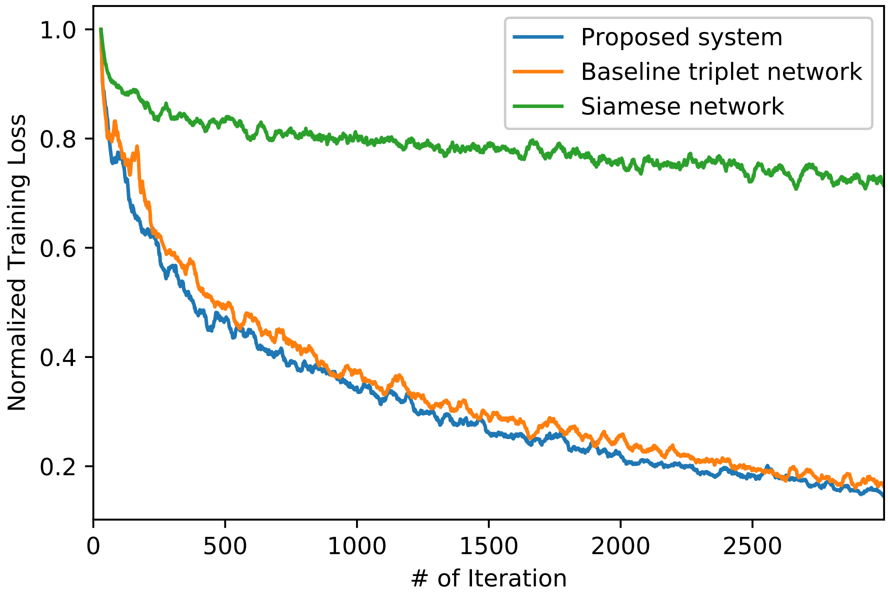

Figure 5.2: normalized training loss trends

5.3.4 2D Visualization of the Feature Vectors

To visualize the proposed system's feature extraction performance, the feature space was projecgted onto two-dimensional Euclidean space using PCA (Figs. 5.6, 5.4, and 5.5). each system's feature extraction performance by determining whether proximate points shared the same identity. Only the first 10 identities were visualized to facilitate the visual confirmation process. The proposed method used the space most efficiently. The Siamese network linearly arranged feature vectors, making it difficult to distinguish between classes. The baseline triplet network caused many identities to overlap.

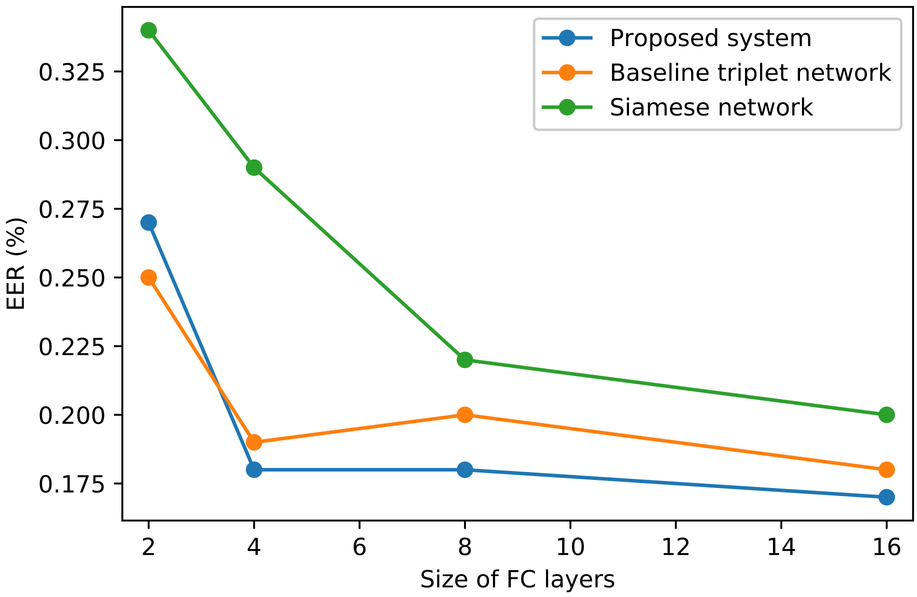

Figure 5.3: Feature Vector Size Effect.

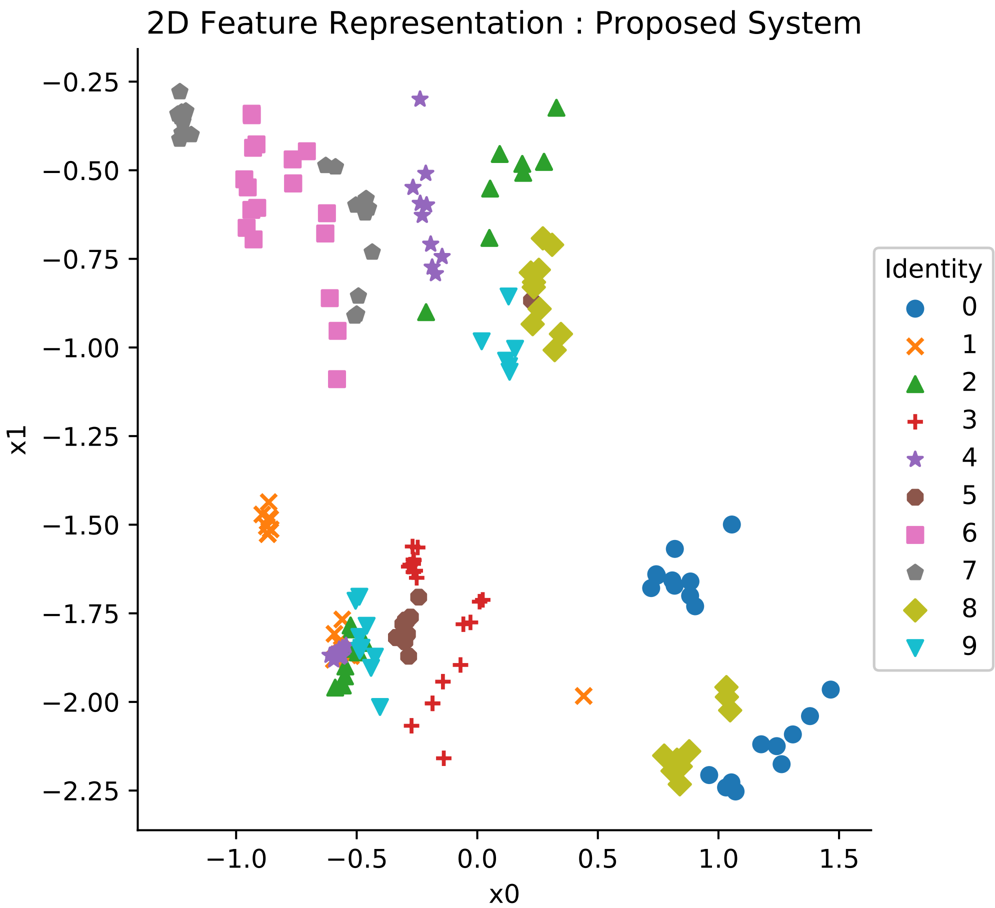

Figure 5.4: 2D Feature Representation : Proposed System.

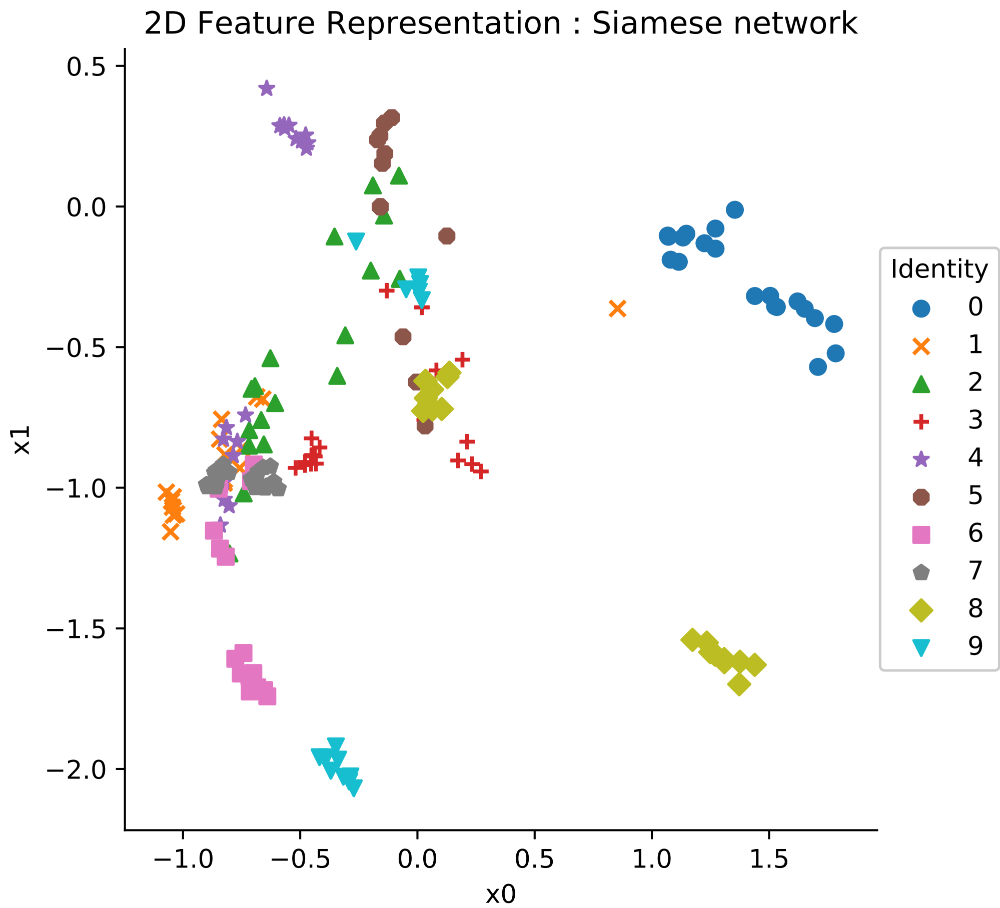

Figure 5.5: 2D Feature Representation : Baseline triplet network.

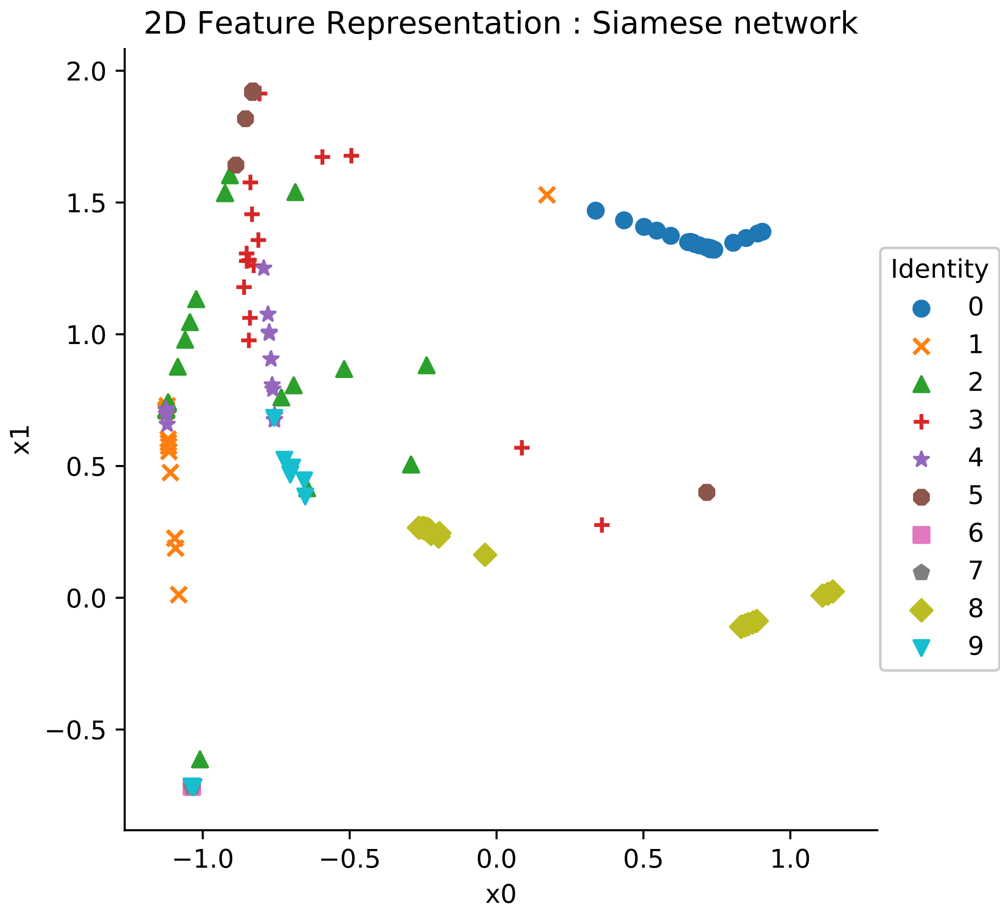

Figure 5.6: 2D Feature Representation : Siamese network.

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] Maged MM Fahmy. Online handwritten signature verification system based on dwt features extraction and neural network classification. *Ain Shams Engineering Journal*, 1(1):59–70, 2010.

[2] Javier Galbally, Moises Diaz-Cabrera, Miguel A Ferrer, Marta Gomez-Barrero, Aythami Morales, and Julian Fierrez. On-line signature recognition through the combination of real dynamic data and synthetically generated static data. *Pattern Recognition*, 48(9):2921–2934, 2015.

[3] Ahmad Sanmorino and Setiadi Yazid. A survey for handwritten signature verification. In *the 2nd International Conference on Uncertainty Reasoning and Knowledge Engineering*, pages 54–57. IEEE, 2012.

[4] Enric Sesa-Nogueras, Marcos Faundez-Zanuy, and Jiří Mekyska. An information analysis of in-air and on-surface trajectories in online handwriting. *Cognitive Computation*, 4(2):195–205, 2012.

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

[6] Gonzalo Bailador, Carmen Sanchez-Avila, Javier Guerra-Casanova, and Alberto de Santos Sierra. Analysis of pattern recognition techniques for in-air signature biometrics. *Pattern Recognition*, 44(10-11):2468–2478, 2011.

[7] Je-Hyoung Jeon, Beom-Seok Oh, and Kar-Ann Toh. A system for hand gesture based signature recognition. In *the 12th International Conference on Control Automation Robotics & Vision (ICARCV)*, pages 171–175. IEEE, 2012.

[8] Hamed Ketabdar, Peyman Moghadam, Babak Naderi, and Mehran 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*, pages 185–188. ACM, 2012.

[9] Jameel Malik, Ahmed Elhayek, Sheraz Ahmed, Faisal Shafait, Muhammad Malik, and Didier Stricker. 3DAirSig: A framework for enabling in-air signatures using a multi-modal depth sensor. *Sensors*, 18(11):3872, 2018.

[10] Heba Abdelnasser, Moustafa Youssef, and Khaled A Harras. Wigest: A ubiquitous wifi-based gesture recognition system. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 1472–1480. IEEE, 2015.

[11] Rajalakshmi Nandakumar, Bryce Kellogg, and Shyamnath Gollakota. Wi-fi gesture recognition on existing devices. *arXiv preprint arXiv:1411.5394*, 2014.

[12] Han-Cheol Moon, Se-In Jang, Kangrok Oh, and Kar-Ann Toh. An in-air signature verification system using Wi-Fi signals. In *Proceedings of the 4th International Conference on Biomedical and Bioinformatics Engineering*, pages 133–138. ACM, 2017.

[13] Akarsh Pokkunuru, Kalvik Jakkala, Arupjyoti Bhuyan, Pu Wang, and Zhi Sun. Neuralwave: Gait-based user identification through commodity WiFi and deep learning. In *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*, pages 758–765. IEEE, 2018.

[14] Cong Shi, Jian Liu, Hongbo Liu, and Yingying 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*, page 5. ACM, 2017.

[15] Kar-Ann Toh, Zhiping Lin, Zhengguo Li, Beomseok Oh, and Lei Sun. Gradient-free learning based on the kernel and the range space. *arXiv preprint arXiv:1810.11581*, 2018.

[16] Liu Yang and Rong Jin. Distance metric learning: A comprehensive survey. *Michigan State Universiy*, 2(2):4, 2006.

[17] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. Deep metric learning for person re-identification. In *2014 22nd International Conference on Pattern Recognition*, pages 34–39. IEEE, 2014.

[18] Kilian Q Weinberger, John Blitzer, and Lawrence K Saul. Distance metric learning for large margin nearest neighbor classification. In *Advances in neural information processing systems*, pages 1473–1480, 2006.

[19] De Cheng, Yihong Gong, Sanping Zhou, Jinjun Wang, and Nanning 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*, pages 1335–1344, 2016.

[20] Shengyong Ding, Liang Lin, Guangrun Wang, and Hongyang Chao. Deep feature learning with relative distance comparison for person re-identification. *Pattern Recognition*, 48(10):2993–3003, 2015.

[21] Faqiang Wang, Wangmeng Zuo, Liang Lin, David Zhang, and Lei 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*, pages 1288–1296, 2016.

[22] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 815–823, 2015.

[23] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop*, volume 2, 2015.

[24] Laura Leal-Taixe, Cristian Canton-Ferrer, and Konrad Schindler. Learning by tracking: Siamese cnn for robust target association. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2016.

[25] Prerana Mukherjee, Brejesh Lall, and Snehith Lattupally. Object cosegmentation using deep siamese network. *arXiv preprint arXiv:1803.02555*, 2018.

[26] Saket Maheshwary and Hemant Misra. Matching resumes to jobs via deep siamese network. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 87–88. International World Wide Web Conferences Steering Committee, 2018.

[27] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[28] Kar-Ann Toh. Analytic network learning. *arXiv preprint arXiv:1811.08227*, November, 2018.

[29] Kar-Ann Toh. Kernel and range approach to analytic network learning. *International Journal of Networked and Distributed Computing*, 7(1):20–28, December 2018.

[30] Kar-Ann Toh. Learning from the kernel and the range space. In *the IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, pages 1–6. IEEE, 2018.

[31] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking*, pages 65–76. ACM, 2015.

[32] Elad Hoffer and Nir Ailon. Deep metric learning using triplet network. In *International Workshop on Similarity-Based Pattern Recognition*, pages 84–92. Springer, 2015.

[33] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

[34] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.

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

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

[37] Kar-Ann Toh and How-Lung Eng. Between classification-error approximation and weighted least-squares learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):658–669, 2008.

국문요약

와이파이 신호를 이용한 공중 서명 인식 시스템

…

핵심되는 말: …