

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

September 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

..

April 2018

**Han-Cheol Moon**

# Contents

<b>List of Figures</b>	<b>4</b>
<b>List of Tables</b>	<b>5</b>
<b>Abstract</b>	<b>6</b>
<b>1 Introduction</b>	<b>7</b>
1.1 Background . . . . .	7
1.2 Motivation and Contributions . . . . .	7
1.3 Organization of Thesis . . . . .	8
<b>2 Literature Survey</b>	<b>9</b>
2.1 In-Air Signature Recognition . . . . .	9
2.2 Wi-Fi Signal based Human Activity Recognition . . . . .	9
2.2.1 Wi-Fi based Human Gesture Recognition Works . . . . .	9
2.2.2 Wi-Fi based Person Identification Works . . . . .	9
<b>3 Preliminaries</b>	<b>10</b>
3.1 Biometric Identification System . . . . .	10
3.2 Evaluation metrics . . . . .	10
3.3 Received Signal Strength Indicator . . . . .	10
3.4 Channel State Information . . . . .	10

3.4.1	Multiple-Input and Multiple-Output . . . . .	10
3.5	Shift Matrix . . . . .	11
3.6	Metric Learning . . . . .	11
<b>4</b>	<b>Proposed System</b>	<b>12</b>
4.1	Overview of the Proposed System . . . . .	12
4.2	Preprocessing . . . . .	12
4.3	Proposed Methodology . . . . .	12
4.3.1	Triplet loss . . . . .	13
4.3.2	Triplet Mining based on the Kernel and the Range space learning . . . . .	15
4.3.3	The ConvNet structures . . . . .	16
<b>5</b>	<b>Experiments</b>	<b>18</b>
5.1	Data set . . . . .	18
5.2	Experimental Settings . . . . .	18
5.2.1	Evaluation Scenarios . . . . .	18
5.2.2	Evaluation Protocols . . . . .	18
5.2.3	Parameter Settings . . . . .	18
5.3	Experimental Results . . . . .	19
5.3.1	Effect of the DC Component . . . . .	19
5.3.2	Effect of the Weighting Scheme . . . . .	19
5.3.3	Effect of the Facing Direction and the Position of the Subject	19
5.3.4	Comparison with the ConvNet . . . . .	19

<b>6</b>	<b>Conclusions and Future Works</b>	<b>20</b>
6.1	Conclusions . . . . .	20
6.2	Future Works . . . . .	20
	<b>Bibliography</b>	<b>20</b>
	<b>Summary (in Korean)</b>	<b>22</b>



# List of Figures

4.1	An overview of the proposed methodology. . . . .	13
4.2	ConvNet structure. . . . .	17

## List of Tables

# Abstract

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

Han-Cheol Moon

School of Electrical and Electronic Engineering

The Graduate School

Yonsei University

...

---

**Key words : Biometrics, Identification, In-Air Signature, Channel  
State Information, and Wi-Fi**

# Chapter 1

## Introduction

### 1.1 Background

...

### 1.2 Motivation and Contributions

... ..

## 1.3 Organization of Thesis

..

## Chapter 2

# Literature Survey

### 2.1 In-Air Signature Recognition

...

### 2.2 Wi-Fi Signal based Human Activity Recognition

.. sec:identification.

#### 2.2.1 Wi-Fi based Human Gesture Recognition Works

..

#### 2.2.2 Wi-Fi based Person Identification Works

..

## Chapter 3

# Preliminaries

..

### 3.1 Biometric Identification System

..

### 3.2 Evaluation metrics

..

### 3.3 Received Signal Strength Indicator

..

### 3.4 Channel State Information

..

#### 3.4.1 Multiple-Input and Multiple-Output

..

### **3.5 Shift Matrix**

..

### **3.6 Metric Learning**



## Chapter 4

# Proposed System

### 4.1 Overview of the Proposed System

.. [1]

### 4.2 Preprocessing

### 4.3 Proposed Methodology

In this paper, we propose person identity verification system based on the in-air handwritten 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[] which utilizes ConvNet[] as a feature extractor. Moreover, to achieve a faster loss convergence, we adopt the kernel and the range (KAR) space learning[] for mining the triplet input. Figure 4.1 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

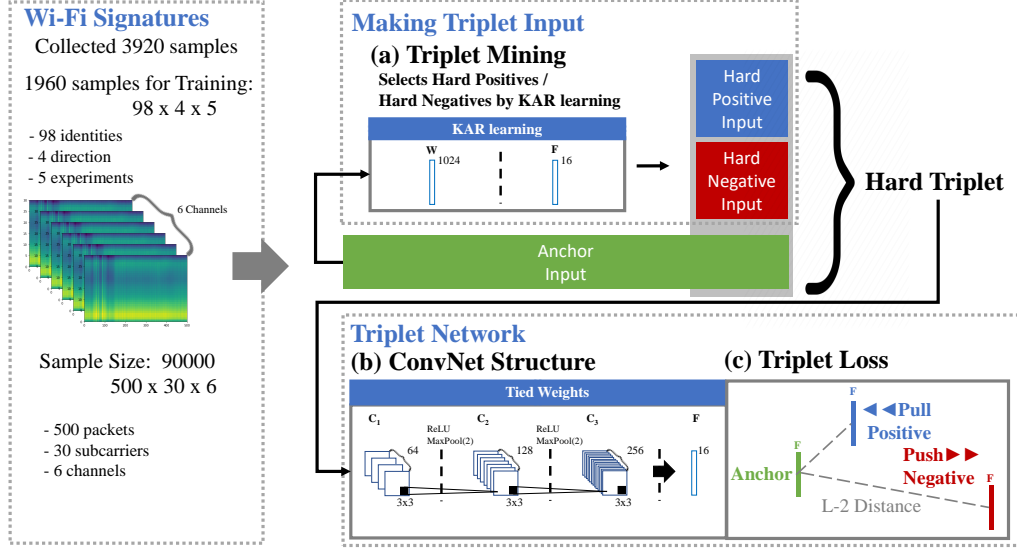


Figure 4.1: An overview of the proposed methodology.

to which likely to be misclassified by the triplet network for 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.

#### 4.3.1 Triplet loss

The purpose of the triplet loss [1] is training the ConvNet structure to learn discriminative features that allocates 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 same identity with that of the anchor while 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  $\alpha$ . Distance measurement function is shown below:

$$dist\{f(x_0), f(x_-)\} - dist\{f(x_0), f(x_+)\} \geq \alpha \quad (4.1)$$

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

$$triplet\_loss = \sum_i^N \max \left( \left[ \|f(x_0) - f(x_+)\|_2^2 - \|f(x_0) - f(x_-)\|_2^2 + \alpha \right], 0 \right), \quad (4.2)$$

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.

### 4.3.2 Triplet Mining based on the Kernel and the Range space learning

In order 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 samples 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 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.

To make the hard triplet input before training 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 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:

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

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 satisfies this property:

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

$$\|KAR(\mathbf{x}_0) - KAR(\mathbf{x}_-)\|_2^2 \leq t_-, \quad (4.5)$$

If the hardest sample, it is known that the outlier data is likely to be selected and there is a risk of overfishing[]. To avoid this problem, 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.

### 4.3.3 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 are made of three layers of ConvNet with triplet loss. Each layers 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)

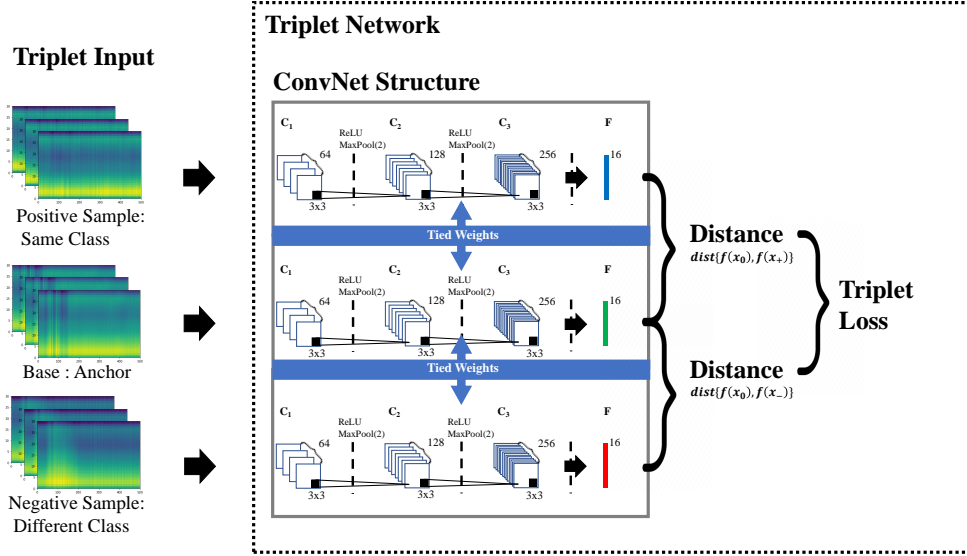


Figure 4.2: ConvNet structure.

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.

# Chapter 5

## Experiments

..

### 5.1 Data set

..

### 5.2 Experimental Settings

..

#### 5.2.1 Evaluation Scenarios

..

#### 5.2.2 Evaluation Protocols

..

#### 5.2.3 Parameter Settings

..

## 5.3 Experimental Results

..

### 5.3.1 Effect of the DC Component

..

### 5.3.2 Effect of the Weighting Scheme

..

### 5.3.3 Effect of the Facing Direction and the Position of the Subject

..

### 5.3.4 Comparison with the ConvNet

..



## Chapter 6

# Conclusions and Future Works

### 6.1 Conclusions

...

### 6.2 Future Works

...

1. Recognizing the in-air signature written along non-LOS.
2. Repeating the same data acquisition experiments in a different place
3. Studying identification performances when multiple people are staying in the same room.
4. To secure a reliable identification system, spoofing attack scenario should be investigated.

# Bibliography

- [1] 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.

국문 요약

와이파이 신호를 이용한 공중 서명 인식 시스템

...

---

핵심되는 말: ...