User Data Selection using CNN Feature Extractor for Fingerprint Localization

Yohei KONISHI†, ***Student member***, **Satoru AIKAWA†a)**, ***Member***, **Shinichiro YAMAMOTO†**, ***Member*** ***,and*** **Yuta SAKAI†**, ***member***

SUMMARY This paper scopes a method that applies CNN to Fingerprint indoor localization. AP information are used to train the CNN. As the number of AP information with correct labels increases, the estimation accuracy by the CNN improves. However, it costs a lot to collect AP information with correct labels. UD (User Data) can be used to solve the problem. The UD is unlabeled data because the measuring method of the UD does not know the user's place exactly. We perform semi-supervise assuming the estimation result of the UD as the correct label. However, the estimation result of the UD may be incorrect. Therefore, it is needed to select UD which is correctly estimated and use it for training the CNN. In this study, we propose a way to select UD by feature value using CNN feature extractor.
keywords: *Fingerprint, indoor localization, CNN, semi-supervised learning*

1. Introduction

Recently, importance of location information has been increasing. Navigation systems using mobile devices such as smartphones have become widespread. The GPS (Global Positioning System) method cannot estimate the exact location in indoor facilities or underground malls because the signal from the satellite is blocked.

A Fingerprint localization based on RSSI (Received Signal Strength Indicator) of Wi-Fi has been proposed as an indoor location estimation. Machine Learning methods especially CNN (Convolutional Neural Network) are applied for the Fingerprint with high accuracy [1,2,3]. AP information with correct labels are used as training data for the CNN. *AP information* are AP identifier and RSSI of observed AP at the coordinates on the map. In general, the more training data used with correct labels for Machine Learning, the better the estimation accuracy can be realized. However, the measurement of a lot of AP information with correct labels is costly. UD (AP information user measured) can be used to solve the problem [4,5]. The UD are unlabeled data because the measuring method of UD does not know the user's place exactly.

We perform a semi-supervised learning assuming the estimated result as the correct label. However, the estimated result may be incorrect. Therefore, it is needed to select UD that estimated correctly and use it for CNN training. In this study, we propose a UD selection method for CNN training. A trained CNN is used for a feature extractor calculation. UD are selected using the feature value extracted from the CNN.

†The author is with University of Hyogo
a) E-mail: aikawa@eng.u-hyogo.ac.jp

In this paper, we confirmed that semi-supervised learning using the proposed method. The effect is verified from the data measured in a building. The proposal can improve the estimation accuracy compared to the CNN trained only using AP information with correct labels.

2. Indoor localization using Fingerprint method

2.1 Fingerprint

This section describes the Fingerprint method for an indoor localization. First, the coordinates are arbitrarily set in the location estimation area. Hereinafter, the coordinates are referred to as *preset coordinates*. The next step is to measure the AP information that is called *preset-AP information*. A *DB*(*database*) is created with the preset-AP information. Here, the preset-AP information are the AP information measured at the *preset coordinates*. A user's terminal measures AP information as *UD* (*User Data*) and compare DB and UD to estimate its location. The most similar coordinate will be the estimated result. The UD is used to estimate the user's location using the CNN and is also used to create a new database to train the CNN for the future.

2.2 Indoor localization method using Deep Learning

This section explains how to apply the Neural Network to the Fingerprint localization. First, preset-AP information are measured by a service provider as shown in section 2.1. Then, the preset-AP information are used as data training the Neural Network. The number of neurons in the input layer is the number of observed AP as shown in Fig 1. The RSSI observed from each AP is input to the first layer. The number of neurons of the output layer is the number of preset coordinates. The final layer with Softmax function outputs the user's existence probability of each coordinate. Ultimately the coordinate with the highest probability is selected as the estimated result.

In this study, we use a CNN for indoor localization. CNN, a type of Deep Learning, repeatedly performs convolutional and pooling operations in the middle layers. The CNN is mainly used for image classification. In this paper, input images are created from RSSI and used to train the CNN[6]. First, the observed AP are arranged on a 2D image. The RSSI values obtained from the AP are considered as pixel values to create the input image for the CNN. The created image is used as data training the CNN.

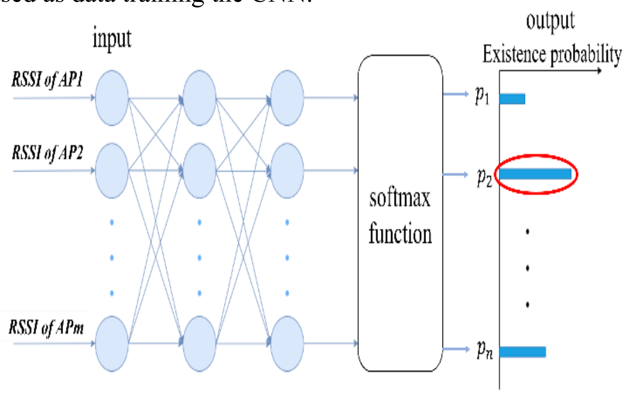


Fig. 1 The flow of indoor localization using deep learning

2.3 Semi-supervised learning using UD

The CNN based localization method using the more data with correct labels, the better the estimation accuracy. However, measuring a lot of preset-AP information is costly for the service provider. On the other hand, it is possible to collect a lot of UD if there are enough users. The UD can be used to train CNN after a user's localization. The cost of measuring preset-AP information can be reduced with the help of the UD. A CNN trained with the preset-AP information is used to estimate the user's location. The estimated result is assumed to be the correct label of the UD. The proposed method can perform semi-supervised learning using the preset-AP information and the UD. Semi-supervised learning is a method of training with both labeled and unlabeled data. However, the labels assumed for the UD may be wrong. Therefore, the estimation accuracy of the CNN may not be improved if all the UD are used. The preset-AP information is measured at the preset coordinate. On the other hand, UD is measured at any position where the user exists. In this reason, if the user is far from the preset coordinates, the estimated results are more likely to be wrong.

Therefore, it is needed to select UD that are correctly estimated. The semi-supervised learning is performed using the *selected UD*. Hereinafter, semi-supervised learning using the selected UD is referred to as *UD-learning*.

3. Proposed method

We propose a method to select UD using extracted feature values from CNN.

3.1 Feature value extracted from CNN

A feature value of input images is extracted from a CNN. A 2D image is input to the trained CNN. The output values are extracted from each neuron in the middle layer as the feature value of the input image [7]. If the two input images are similar, their feature values are close. In a proposed scheme, UD is selected if the feature values of UD and preset-AP information are close.

The configuration of the CNN is shown in Fig.2. The output values in the middle layer before the output layer of the CNN were used as feature value to select the UD.

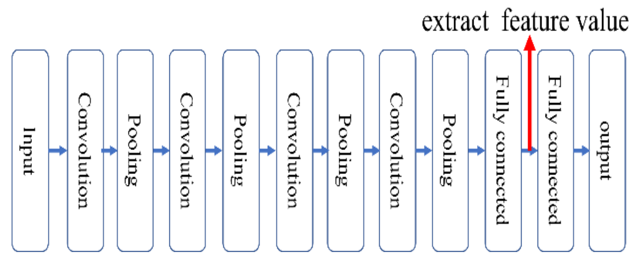
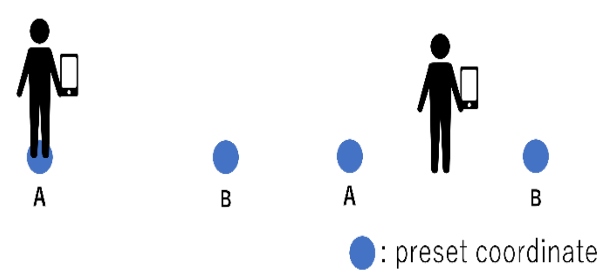


Fig. 2 Extract feature value using CNN

3.2 UD selection method

The UD that should be added to the CNN training data is the one measured near the preset coordinates as shown in Fig.3(a). In contrast, if the measured position is far from estimated coordinates, such as the center between preset coordinates as shown in Fig. 3(b), the UD should not be added to the training data. The feature value of the UD measured near the preset coordinate is similar to the feature value extracted from the preset-AP information. Contrary to this, the feature value of the UD measured away from the preset coordinates is different. Therefore, the feature values extracted are used as an indicator to classify their distance.



(a) UD that should be used (b) UD that should not be used
Fig. 3 UD used for CNN training

First, preset-AP information are measured as described in Section 2.2. The CNN is trained using images created from the preset-AP information as described in Section 2.3. After the CNN trained, the feature values are extracted from all the training data. Then, the extracted feature values of all the preset coordinates and their average values are gathered as shown in Fig. 4. In the table of Fig. 4, $1, 2, 3, \cdots, n$ show the outputs of the neurons in the middle layer, and (1), (2), (3) $\cdots$ show the number of preset-AP information measured at the same preset coordinate. The average value is used as the feature value for each preset coordinate.

A coordinate is selected by the location estimation for the UD using the CNN. At that time, the feature value is extracted of the UD from the CNN. Then, the feature value of UD is compared with that of the preset coordinate selected by the location estimation. If the two feature values are similar, they are used as training data for the CNN. In this study, the Euclidean distance, shown in the following equation, is used to calculate the similarity between two feature values. $\boldsymbol{x}, \boldsymbol{y}$ are the feature values, d is the Euclidean distance. The smaller the value of Euclidean distance, the more similar the two features are.

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \qquad (1)$$

Coordinate : A

	1	2	3	・・・	$n-1$	n
(1)	0.7	1.8	-0.5	・・・	1.7	-2.4
(2)	0.4	1.6	0.1	・・・	2.1	-1.8
(3)	0.3	2.2	-0.7	・・・	1.4	-1.9
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Calculate the average value

Feature value	0.5	1.9	-0.4	・・・	1.8	-1.7

Fig. 4 Feature value creation at each coordinate

4. Verification

4.1 Verification environment

The verification environment is shown in Table1 and Fig.5. In this verification, the preset-AP information are measured 40 times for each preset coordinate to train the CNN as shown in Fig. 5. We measured 100 UD at each *UD coordinate* where the user exists. The UD coordinates are 0.25m apart from each other as shown in Fig.5. The 80 UD of the 100 UD measured were used for UD-learning, and the 20 were used as test data to verify the estimation accuracy of the CNN created by the UD-learning. The UD selected using the Euclidean distance are added to new training data. Then, the effect of the UD-learning is verified by the estimation accuracy of the CNN.

Table 1 Verification environment

Measurement place	University building corridor
The number of coordinates	preset coordinates : 11 UD coordinates : 121
Coordinate interval	preset : 3m UD : 0.25m
The number of measurements	40 for each preset coordinate 100 for each UD coordinate

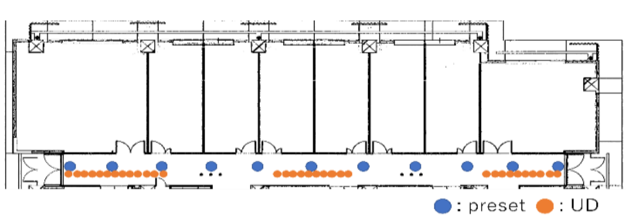


Fig. 5 The verification environment figure

4.2 Verification of UD selection

First, the CNN were trained using the preset-AP information in this verification. Next, location estimation was performed for the UD using the trained CNN. Then, the Euclidean distance that means the similarity was calculated from the extracted UD feature value using. Fig. 6 shows a histogram of the calculated Euclidean distance. Fig. 6 shows the error between the estimated position and the actual measured position. Estimation error for UD is less than 1.5 m, which is half of the preset coordinate interval of 3 m, is shown in blue, and estimation error for UD is more than 1.5m is shown in red. Fig. 6 shows that as the Euclidean distance decreases, the amount of data with small estimation error increases. Therefore, the UD with smaller Euclidean distance are appropriate to the training data for CNN.

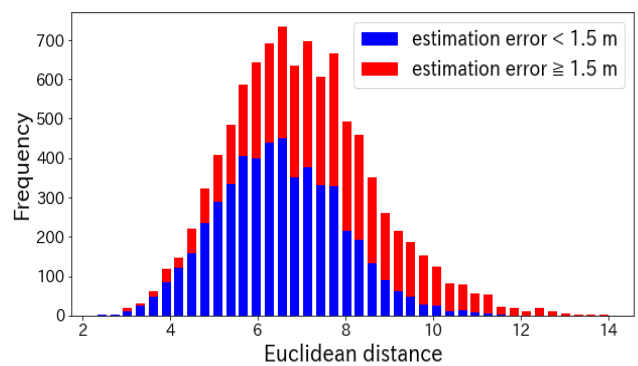


Fig. 6 Histogram of Euclidean distance and estimation error

4.3 Estimation accuracy of CNN trained using UD

A UD-learning was performed in the above verification environment. The location was estimated using the trained CNN by UD-learning. Fig. 7 shows the average error with respect to the threshold. In Fig. 7, "*before update*" shows a CNN trained only with preset-AP information. Fig. 7 shows that the UD-learning improves the estimation accuracy over "before update". In this verification, the average error was minimal at the threshold set to 6.5, where average error was 1.75m.

Since the average error was 2.11m before update, the UD-learning reduced the average error by up to 36 cm. When the threshold is set larger, the number of selected UD increases. However, the value of the average error is not significantly smaller than before update. The reason is probably that by setting the large threshold, a large number of UD far from the preset coordinates are used for UD-learning. These results show the effectiveness of the UD selection method using the feature value extracted.

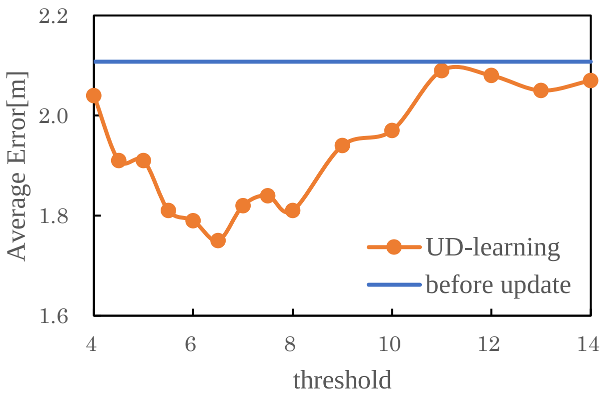


Fig. 7 Estimation accuracy with changing threshold

Fig. 8 shows the cumulative relative frequency of the estimation error for the test data with a threshold of 6.5. Fig.8 shows that the estimation error is smaller than before update. Therefore, the UD-learning can improve the estimation accuracy of CNN with an appropriate threshold.

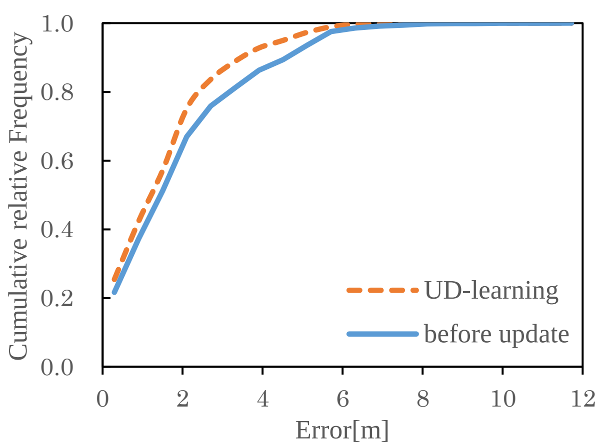


Fig. 8 Cumulative relative frequency of semi-supervised learning using UD

5. Conclusion

In this study, we proposed a method to select UD for the semi-supervised learning using a feature extractor from CNN. Euclidean distance was used to calculate the similarity of feature values between preset-AP information and selected UD. The UD-learning can improve the estimation accuracy adding UD whose similarity is less than the appropriate threshold. The effectiveness of the proposed UD selection method was confirmed from the experimental results.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 21K04065.

References

[1] Satoru Aikawa, et al., "WLAN Finger Print Localization using Deep Learning," IEEE APCAP Aug. 2018. DOI:10.1109/ APCAP. 2018. 8538306.
[2] Linchen Xiao, et al., "A Deep Learning Approach to Fingerprinting Indoor Localization Solutions", International Telecommunication Networks and Applications Conference (ITNAC), Nov. 2017 DOI: 10.1109/ATNAC.2017.8215428.
[3] Ayush Mittal et al., "Adapting Convolutional Neural Networks for Indoor Localization with Smart Mobile Devices," GLSVLSI'18, Chicago, IL, USA, May 23-25, 2018.
[4] Chunjing Song and Jian Wang, "WLAN fingerprint indoor positioning strategy based on implicit crowdsourcing and semi-supervised learning," *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 11, 356, 2017.
[5] Yusuke Miyamoto, et al., "User data selection scheme to data-base update errors for fingerprint localization,", IEICE Communication Express, vol.10, no5(2021), 343–348.
[6] Satoru Aikawa, et al., " CNN Localization Using AP Inverse Position Estimation" 2019 IEEE Conference Antenna Measurements and Applications(CAMA2019), OCTOBER23-25,2019. DOI:10.1109/CAMA47423.2019.8959663.
[7] Hideki Nakayama, "Image feature extraction and transfer learning using deep convolutional neural networks ", IEICE Technical Report, vol. 115, no. 146, SP2015-45, pp. 55-59, July.2015 (In Japanese).