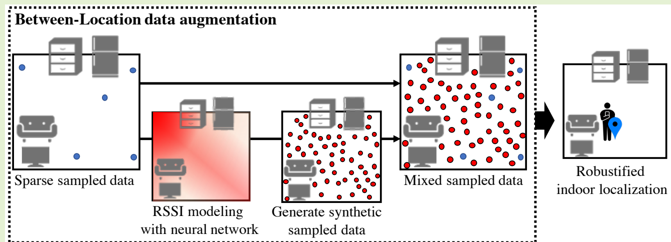
Robustifying Wi-Fi localization by Between-Location data augmentation

Masato Sugasaki, Masamichi Shimosaka

***Abstract*— Wi-Fi fingerprint-based indoor localization is one of the most practical localization methods, which does not require extra infrastructure and special hardware. However, we need to acquire a dataset with a high-density dataset in the target environment in this framework. To overcome the data acquisition cost problem, we propose a brand new data augmentation for Wi-Fi indoor localization named Between-Location data augmentation (BL data augmentation). We generate the fingerprint data for the whole target environment with high density by only using the sparsely sampled data. Between-Class learning, which is the origin of BL data augmentation and the latest powerful data augmentation method for sound recognition and image processing, mixes two data linearly with normalization; however, this mixing does not make sense in indoor localization because mixed fingerprint has no meaning and the label of indoor localization is not categorical information but physically correlated information. To overcome these two problems, we propose the generative model based on neural networks installed the physical relationship of labels and Wi-Fi fingerprint property. BL data augmentation enables us to reduce data sampled locations while keeping the localization accuracy even if some target locations have no data. From the experimental results, indoor localization methods with BL data augmentation outperform the state-of-the-art data augmentation method on several indoor localization models, whatever the data collection location is dense or sparse. Moreover, the localization with BL data augmentation using 10 % sampled location achieves the same accuracy with localization without data augmentation using all sampled locations.**



***Index Terms*— Data augmentation, Fingerprinting, Indoor positioning, RSSI modeling, Virtual sensing**

I. Introduction

Wi-Fi-based indoor localization [1], [2] is one of the most practical techniques because of performing only with Wi-Fi access points (AP) already installed in the target environment. In the Wi-Fi indoor localization methods, the Wi-Fi fingerprint-based indoor localization [1], [3] is the most attractive method due to the availability of most of all devices without installing special hardware or software. In this framework, a fingerprint that uses the vector of signal strength of Wi-Fi called RSSI is used as an input of the localization model.

However, we should acquire the fingerprint data in the target environment with high density, and this labor-intensive cost prevents us from practical use. To reduce the data acquisition cost, some work tries to acquire the data with less effort by crowd-sensing or unsupervised training; however, the number of locations to acquire the data is not decreased, and the acquisition cost is not decreased fundamentally. These methods distribute the data acquisition cost to the consumers or crowd workers. However, these methods cause some other costs such as budget cost or waiting time for enough data acquisition. These other costs also prevent indoor localization from practical use even if the data acquisition cost is resolved. To reduce the data acquisition cost fundamentally, we need to reduce the data acquisition location, while keeping the localization accuracy in the whole environment even if some target locations have no data. From the machine learning perspective, this problem is how to improve the estimation performance for unseen data, which is known as a difficult setting named model generalization problem or zero-shot learning. In other domains, especially on the image processing, the data augmentation method [4], [5] is actively explored as a possible solution for model generalization problems. The data augmentation method generates the variation of data to increase the generalization performance or avoiding overfitting.

Recently, a powerful data augmentation method named Between-class learning (BC learning) was proposed for sound recognition and image processing [6], [7]. BC learning generates the synthetic data by mixing two data on different classes with a random ratio. The advantage of BC learning not only improves the generalization performance but also installs the class relationship by the synthetic data between two classes.

By applying the BC learning for indoor localization to generate unseen data (i.e., data for not sampled location), the accurate localization model seems to be constructed with sparse sampling location; however, there are two problems to apply: 1) mixing the Wi-Fi signal strength does not make sense and 2) label of the indoor localization is not categorical information but a location that has a physical relationship between each label. In sound recognition or image processing, the mixed data becomes sound data or image data, and a mixed label indicates the two categories are mixed. However, in the fingerprint, the directly mixed fingerprint is just the average of signal strengths not related to any location. Therefore, it is inappropriate to just apply the BC learning for the Wi-Fi indoor localization.

Thus, we propose the new data augmentation method for indoor localization based on BC learning named Between-Location data augmentation (BL data augmentation). We propose the deep model-based fingerprint data generation at not data sampled location by focusing on the relationship of the location in each sampled data. We construct the generative model for Wi-Fi signals of each AP by using sparse data, then we form the synthetic fingerprint dataset that virtually sampled the whole of the target environment with high density. By mixing the generated synthetic dataset and sampled dataset, we achieve a more robust localization model whatever the data collection location is dense or sparse. Moreover, our BL data augmentation can be applied to most indoor localization methods regardless of the model type because our model just requires mixing the generated synthetic data to the acquired dataset. The contributions of our paper are as follows:

- We proposed a brand new data augmentation method for indoor localization that installs the location relationship and Wi-Fi signal property.
- We provide how to improve the performance of the indoor localization model with sparsely sampled locations regardless of the model type includes the linear model and deep-structure model.
- We evaluated BL data augmentation with an actual dataset on neural-based and linear indoor localization models and demonstrated that our proposed model improves the performance in any type of localization method whatever the data collection location is dense or sparse.

Related work:

Wi-Fi-based indoor localization: Wi-Fi-based indoor localization [1], [8], [9] is one of the most practical methods because the localization is performed only using the Wi-Fi signal of AP installed in the target environment in various indoor localization [10]–[12]. In this framework, Wi-Fi signal information, such as RSSI and channel state information [2], [9], [13], is used to estimate the location of the target device. Among the Wi-Fi indoor localization, fingerprint-based localization is one of the practical techniques because this technique works with almost all devices that support Wi-Fi. Fingerprint-based localization estimates the location from the fingerprint, which is a vector of RSSIs.

However, in fingerprint-based localization, the data acquisition process is problematic because of a time-consuming and labor-intensive task. To construct the robust indoor localization model, we need to acquire the fingerprint data whole the target environment finely, several times. This task prevents the practical use of fingerprint-based indoor localization.

Reducing the data acquisition cost in indoor localization: In the last decade, some researchers try to resolve the data acquisition cost problem with various approaches. Crowd-sensing [14]–[16] and unsupervised learning [17]–[20] have been explored as methods of increasing the data volumes without increasing the provider cost. In the crowd-sensing approach, they ask crowd workers to acquire the data or explore which places are inaccurate. In the unsupervised learning approach, they try to create a localization model using a large dataset uploaded by consumers. In these frameworks, we can distribute the data acquisition cost to crowd workers and reduce the provider cost; however, the data amount required to construct a localization model is not decreased. The acquisition of a large dataset causes large budget costs in the crowd sensing or long waiting period in the unsupervised method.

Semi-supervised learning [21]–[23] have also been explored as an indoor localization method with low effort. The semi-supervised learning methods use the small labeled data and large unlabeled data to construct the more accurate localization model or less unlabeled data than the unsupervised framework. In this framework, the waiting period of one of the significant issues on the unsupervised frameworks is decreased; however, the data acquisition on crowds still needed, and the waiting period have remained.

RSSI distribution modeling [24]–[28] was explored as the other way to reduce the data acquisition cost by modeling the RSSI radio map. In this framework, there is two way to modeling; one is a GPR-based radio map modeling and another one is wave propagation path-loss model-based modeling. In this framework, once we construct the RSSI radio map for each AP in the target environment, we can generate the fingerprint or use the map matching method to localize. However, we need to acquire the dense dataset in the target environment to construct a radio map in the GPR-based method. This data acquisition is problematic same as the fingerprint localization method discussed in the previous section.

The path-loss model-based method requires the physical property of the environment such as the accurate floor map information, AP location information, and furniture placement information or dense dataset to optimize with dense sampled dataset. As for the physical property of the environment, we need to accurate site survey of the physical property and it is time-consuming and labor-intensive as well as dense data acquisition. Moreover, the localization service provider sometimes cannot access the floor map and access point location information due to security issues.

Therefore, these methods cannot fundamentally resolve the data acquisition cost problem. We need to reduce the data sampled locations in the target environment to resolve the data acquisition cost problem fundamentally.

Data augmentation in other domains: In image processing, data augmentation methods [4], [5], [29], [30] are developed to resolve data imbalance and data invariance problem. The data augmentation method generates the new image by rotating, moving, scaling, add noise, and so on. With the data augmentation, the estimation model is generalized by avoiding overfitting, especially for the mapped images (i.e., rotated, moved, and scaled).

Recently, BC learning [6], also known as mixup [7], was proposed as a data augmentation approach for image processing and speech recognition. This framework generates artificial data by mixing real data at a randomly generated ratio; hence,

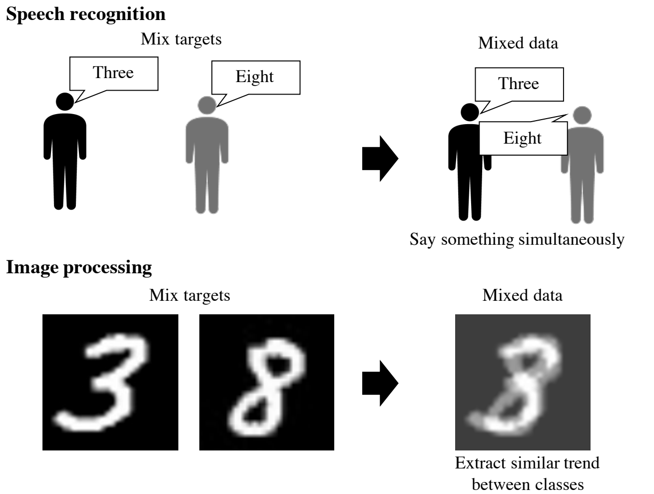


Fig. 1: Mixed data in sound recognition and image processing

we can generate a large amount of artificial data from a small real dataset. BC learning not only generates the synthetic data but also installs the relationship of the class categories thanks to the generated data with two label properties. This property improves furthermore generalization performance. However, BC learning does not make sense at indoor localization. We discuss the data augmentation in BC learning and problems on indoor localization in the following sections.

II. Wi-Fi fingerprint-based indoor localization setting and BC learning reviews

A. Problem setting of Wi-Fi-based indoor localization

A Wi-Fi-based fingerprinting indoor localization model is constructed by the RSSI fingerprint $\boldsymbol{x} \in \mathbb{R}^d$ labeled by location $y \in \mathcal{Y}$, where $d \in \mathbb{N}$ is the number of AP available in the target environment. Let the fingerprint dataset be $\mathcal{D} = \{(\boldsymbol{x}^{(i)}, y^{(i)})\}_{i=1}^{n}$. By using the acquired labeled fingerprint dataset, we construct the indoor localization model that is defined as $f(\boldsymbol{x}) : \mathbb{R}^d \to \mathcal{Y}$ such as the regression model, multiclass-classifier model, and deep learning-based model.

In general, we need to acquire the fingerprint data from dense locations in the target environment to achieve enough accuracy localization model. For example, we want to construct the localization model with $3\ \mathrm{m}$ average error, we need to collect the dataset more precisely such as $1\ \mathrm{m}$ by $1\ \mathrm{m}$. If we collect the data sparsely, the localization accuracy is degraded and we cannot achieve the required accuracy. Thus, we propose the BL data augmentation that generates the synthetic fingerprint data by enhancing the BC learning to the Wi-Fi fingerprint-based indoor localization. BL data augmentation improves the localization performance whatever the data sampled location is dense or sparse.

B. Review of BC learning framework in sound recognition and image processing

Before discussing BL data augmentation, we review BC learning for the sound recognition and image processing that improves the model performance. In BC learning in sound recognition and image processing, the data generation method that matches the data property is a key point for improving the model performance.

In the sound recognition and image processing case, BC learning is easily applicable because of two properties: 1) the mixed data can be simply generated from two input data, and 2) the mixed class can be defined by linear interpolation. In the sound recognition case, we generate the mixed data $x^{(G)}$ from the two sound data with different labels $\{x^{(1)}, l^{(1)}\}$ and$\{x^{(2)}, l^{(2)}\}$ with the posterior probabilities of the mixed data for each mixed label $P(l^{(1);x^{(G)}}) = r$ and $P(l^{(2);x^{(G)}}) = 1 - r$ as $\frac{p\boldsymbol{x}_1+(1-p)\boldsymbol{x}_2}{\sqrt{p^2+(1-p)^2}}$ where $p = \frac{1}{1+10^{\frac{G_1-G_2}{20}}\frac{1-r}{r}}$. It is natural data generation because the sounds can be overlapped, e.g., two people saying ”three” and ”eight” simultaneously, as shown in Fig. 1. In the case of image processing, we generate the mixed data $x^{(G)}$ by regarding the two input data $\{x^{(1)}, l^{(1)}\}$ and$\{x^{(2)}, l^{(2)}\}$ as a wave form data to create natural mixing expression as $\frac{p(\boldsymbol{x}_1-\mu_1)+(1-p)(\boldsymbol{x}_2-\mu_2)}{\sqrt{p^2+(1-p)^2}}$ where $p = \frac{1}{1+\frac{\sigma_1}{\sigma_2}\frac{1-r}{r}}$, μ is the static component when x is represented as $x = \mu + \mathbf{d}$ with the wave component $\mathbf{d}$, and σ is a standard deviation for the image. From this mixing method, we generate effective synthetic image data to extract similar features for two classes for the neural image processing model as shown in Fig. 1.

However, Wi-Fi-based indoor localization does not follow these two properties. We discuss the difficulty of applying BC learning and propose a novel data augmentation method in the next section.

III. Between-Location data augmentation for Wi-Fi indoor localization

In this section, we first discuss the requirement of BC learning in Wi-Fi fingerprint-based indoor localization. Next, we present the overview of the data augmentation method to improve the accuracy of Wi-Fi indoor localization with sparse data sampling, what we called Between-Location data augmentation. We next discuss the property of Wi-Fi signal related to the locations to define the Wi-Fi fingerprint generation model. Finally, we propose the fingerprint data generation model for BL data augmentation.

A. Requirement of BC learning for Wi-Fi indoor localization

Wi-Fi fingerprint-based indoor localization does not follow the two properties of the requirement of applying BC learning introduced in section II-B. As for the first property, we cannot generate the mixed fingerprint data by linear interpolation due to the radio wave propagation and the environmental dependencies. As for the second property, the class in indoor localization indicates the actual location and there is a physical relationship between each class. From the physical relationship, the label for generated data does not indicate mixed label of y_1 and y_2 with the probability $P(y^{(1)}; x^{(G)}) = r$ and $P(y^{(2)}; x^{(G)}) = 1 - r$ but actual different location y_3. Therefore, the naive data mixing method applied to the sound

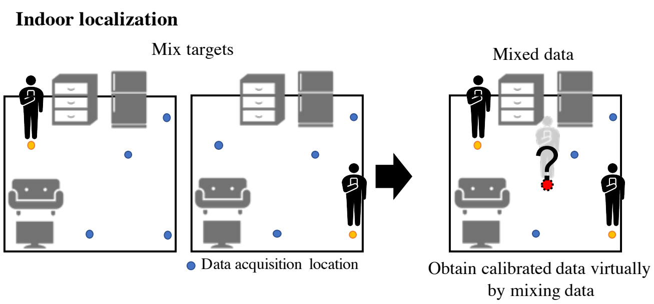


Fig. 2: Mixed data in indoor localization

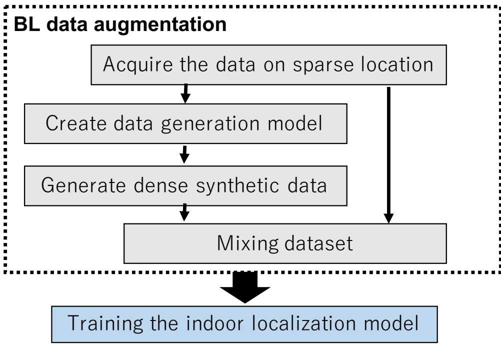
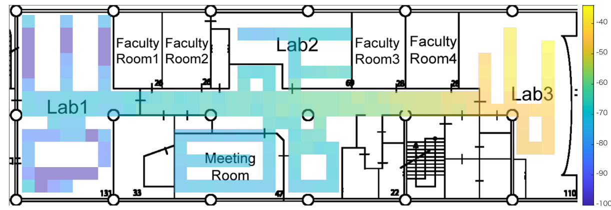
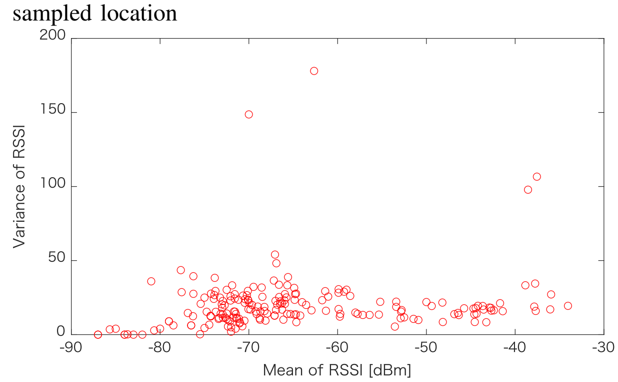


Fig. 3: System overview



(a) Mean RSSI in the target location: not deep blue locations is data sampled location



(b) Relationship between mean and variances in each location

Fig. 4: RSSI characteristics from the view point of the location

recognition and image processing does not make sense in the Wi-Fi fingerprint data.

To enhance the BC learning idea to the Wi-Fi fingerprint, we need to construct the data augmentation method that matches the following condition: 1) considering the Wi-Fi signal property, 2) considering the location relationship on the labels. As for the matching to the Wi-Fi signal data property, the Wi-Fi signal modeling considering the environmental condition is required. As for matching the localization application property, labeling by actual location for the generated synthetic data is needed. To match these requirements, we construct the BL data augmentation to obtain the data augmentation for the Wi-Fi indoor localization that performs as the virtual data sampling in the data acquisition location shown in Fig. 2.

B. Overview of BL data augmentation for indoor localization

We propose a brand new data augmentation method for Wi-Fi indoor localization that matches the two requirements, 1) considering the Wi-Fi signal property, 2) considering the location relationship on the labels. We discuss the Wi-Fi signal property on the real environment in the section III-C and present how to construct the data generation model at section III-D in detail.

Fig. 3 shows the overview of the BL data augmentation framework architecture. In data acquisition, we acquire Wi-Fi fingerprint data from the target environment with the sparse sampling location. Using the acquired dataset, we construct the generative model for the fingerprint data to create synthetic data virtually acquired in the target environment. We generate the synthetic fingerprint dataset with the constructed generative model to the whole target environment densely. Finally, we create the training dataset by mixing the generated synthetic fingerprint dataset and the acquired dataset.

By using the generated dataset by BL data augmentation, we construct the more robust indoor localization model with sparse location sampling. It should be noted that BL data augmentation can be applied to almost all indoor localization methods that include neural network-based regression, multiclass-classification, linear regression, and linear multiclass-classification models. It also should be noted that the BL data augmentation works on the offline phase of the indoor localization that is the model construction part.

C. Relationship between a location and Wi-Fi RSSI

Fig. 4a shows the mean RSSI value on each location and Fig. 4b shows a relationship between the mean and variances in each location in the dataset (that used in the experiment). We observed that the RSSI means are changed with the location, and the variances are around 20 (i.e., the standard deviation is around 4 to 6 dBm) even if the location or mean of RSSI are changed. From this observation, the RSSI distribution at a single location $p(x|y)$ can be approximated to the Gaussian distribution.

Thus, the RSSI model on location y is simply formulated as follows:

$$\hat{x}_{i,y} \sim \mathcal{N}(\bar{x}_{i,y}, \sigma^2), \tag{1}$$

where $\bar{x}_{i,y}$ is the mean of RSSI in the location y for i-th AP and σ^2 is the variance of RSSI. When we estimate the mean value for the location that the data is not sampled, we can generate the synthetic dataset of Wi-Fi fingerprint. In next section, we provide the deep network-based method for estimating $\bar{x}_{i,y}$ (i.e., wave propagation model) from sparse data.

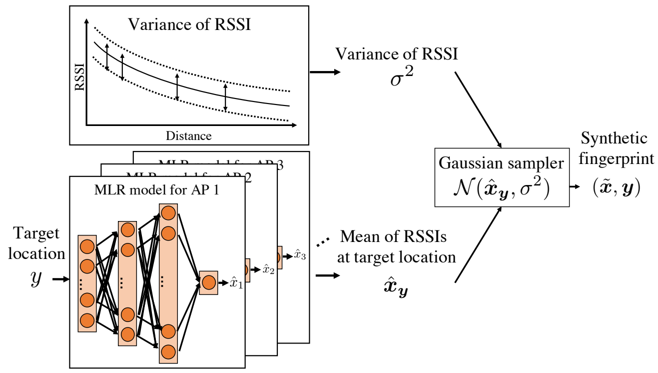


Fig. 5: Synthetic fingerprint generation with MLR and Gaussian sampler

D. BL data augmentation: fingerprint generation model from the location information

To generate the synthetic dataset of Wi-Fi fingerprint, we define the RSSI generation model as follows:

$$\hat{x}_i \sim \mathcal{N}(g_i(y), \sigma^2), \tag{2}$$

where $g_i(y) : \mathcal{Y} \to \mathbb{R}$ is the mapping function of RSSI from the location and σ^2 is the variance of RSSI. We can generate synthetic data by constructing the precise RSSI estimation model even if the target locations have no data.

We choose the multi-layer regression (MLR) model that is a deep learning-based model as modeling of a mean of RSSI in each AP. MLR model has the flexibility to model the non-linear structure. In RSSI, there is a non-linear propagation due to the target environment structure, such as the doors, walls, and furniture. MLR model can install these properties from the dataset even if the sampled location is sparse while keeping the smoothness of RSSI on the near location without applying any radio propagation model explicitly. From this functionality, the MLR model needs a few measurement locations for modeling the wave propagation in the target area or room. Moreover, the MLR model can model the RSSI radio map includes environmental property only with the RSSI data and this model does not require any floor map and AP location information. This smooth modeling and fitting to the sparse data match to consider RSSI propagation and the environmental effects simultaneously.

It should be noted that the Gaussian processes [31], [32] or path-loss model [33] that often used in the RSSI modeling is considered as a possible solution to constructing the $g_i(y)$; however, these models do not match to the sparsely sampled location setting. In the Gaussian processes, the variances on the location without data become large, and this change does not follow the observed property. Because of this variance changing, the quality of the generated data, especially on the location without data becomes low. In the path-loss model, we can fit the parameter from the sparsely sampled location data; however, we cannot reflect the target environment structures because of the fixed parameterized model. In comparison with these models, MLR provides the most proper modeling for the indoor localization model.

We employ the simple MLR model that has four fully connected layers with the hyperbolic tangent activation function. We apply the least squared error $\sqrt{(\hat{x}-x)^2}$ as a loss function of this neural network. We train the MLP model with the acquired dataset on sparse location for d-th AP, respectively, and get the d-th MLP models. From these MLP model, as shown in Fig. 5, the synthetic fingerprint data $\tilde{\boldsymbol{x}}$ is generated on $\boldsymbol{y}$ with Gaussian sampler as follows:

$$\tilde{\boldsymbol{x}} \sim \mathcal{N}(\hat{\boldsymbol{x}}_{\boldsymbol{y}}, \sigma^2), \tag{3}$$

where $\hat{\boldsymbol{x}}_{\boldsymbol{y}} = (g_1(\boldsymbol{y}) \cdots g_d(\boldsymbol{y}))$ is the vector of the estimated mean by trained MLR model for each AP.

We construct the synthetic fingerprint dataset by generating multiple fingerprints to the dense location in the target location. From the generation, we obtain the dense fingerprint dataset for the target location. This synthetic dataset helps that the localization model capture the fingerprint characteristics related to each location.

By using the mixed dataset of the synthetic fingerprint dataset generated by BL data augmentation and the acquired dataset as a training dataset, we construct a more robust localization model than trained only by the acquired dataset. It should be noted that the BL data augmentation method is independent of the localization model, we can apply any indoor localization models even if the model structure is highly complicated such as the deep localization model that recently advanced.

It also should be noted that our model ignores the APs that are not included in the acquired dataset. However, these APs provide weak signals and less contribution for localization than other APs with strong signals. From this fact, our model provides accurate localization even if some weak signal APs are ignored.

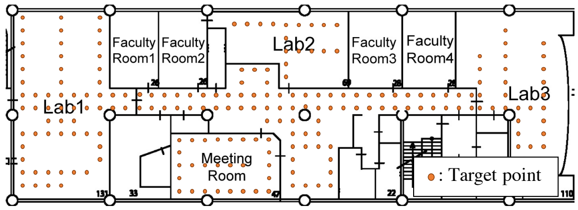
E. Analysis of the data acquisition cost

Here, we discuss the data acquisition cost in the sparsely sampled location dataset. Let $\mathcal{Y}^{(\mathrm{S})} \bigcup_i y^{(i)} \subset \mathcal{Y}$ is the sampled location at the sparse dataset. The reduced ratio of data acquisition location can be formulated as $\frac{|\mathcal{Y}^{(\mathrm{S})}|}{|\mathcal{Y}|}$. This ratio directly affects the data acquisition cost, including the time of data acquisition and the number of fingerprint scanning.

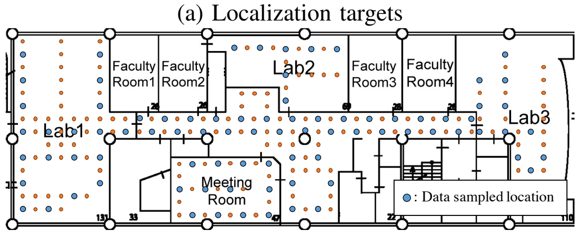
For example, we spent 5 hours acquiring the data at the experimental target environment shown in Fig. 6a (5 times fingerprint scanning at all target locations). When the reduced ratio is 15% shown in Fig. 6c, we only need to acquire 45 minute to acquire the dataset. From this case, it can be said that the data acquisition cost is drastically reduced.

IV. Experiment

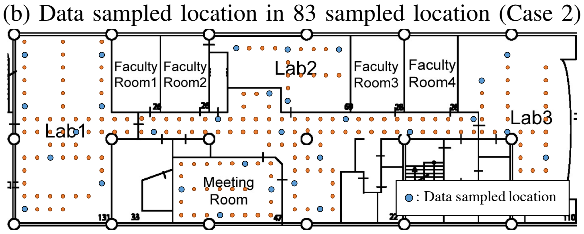
We evaluate the performance of BL data augmentation with two settings: 1) the performance under the different number of the sample locations, and 2) the performance on three different designed sampled locations.



(a) Localization targets



(b) Data sampled location in 83 sampled location (Case 2)



(c) Data sampled location in 30 sampled location (Case 3)

Fig. 6: Target location of the environment

A. Experimental setting

To evaluate the model, we acquire the dataset from a floor of a building of the university that area is $15\ \mathrm{m} \times 40\ \mathrm{m}$ as shown in Fig. 6. We set 210 localization target locations on the floor. We sample the data from each target location with Nexus 5. We use the 5 fingerprint data in the selected sampled locations that depend on each testbed as the training data. As for the test dataset, we use the 30 data for all target locations, 6300 in total. It should be noted that some target location does not have a training data depends on the sparse sampled location setting (described in each experiment part).

B. Comparison methods of data augmentation

In this experiment, we compare the three data augmentation methods on three localization models as follows.

1) Data augmentation methods: We generate 10 data for each target location depicted in Fig. 6a and 2100 data are augmented for the training indoor localization model with our BL data augmentation and the following comparison models.

Without data augmentation (W/O): For the baseline, we employ the localization model without data augmentation.

Linearly interpolation for normal BC Learning (BC-Linear): For evaluating the normal BC Learning performance [6], we employ the linear interpolation with $r\boldsymbol{x}_1 + (1-r)\boldsymbol{x}_2$, where $r = \frac{\Delta(\hat{y},y_2)}{\Delta(\hat{y},y_1)+\Delta(\hat{y},y_2)}$ and $\Delta(y_1, y_2)$ is function to calculate the distance between y_1 and y_2.

Gaussian processes based data generation (BC-GPR): To compare the data generation performance, we prepare the Gaussian processes regression (GPR)-based data generation. We used the Gaussian kernel $\exp(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{\sigma^2})$ as a kernel of GPR and we use 1.0 as parameter σ. We set -100 as a mean of prior distribution and 2.0 as a variance of prior distribution.

It should be noted that we decided hyper-parameters with checking the RSSI generation error on dense sampling location settings for all data augmentation includes a proposed method that is shown in section IV-D.4.

2) Localization models: We employ the three localization methods to evaluate the effect of data augmentations. We selected two classification methods and one regression method to localize as follows:

Linear multiclass classifier with cost-sensitive hinge loss (Linear): We employ the multiclass classifier with cost-sensitive hinge loss that considers the physical distance of each class described in [34] as a linear localization model.

Deep multi-layer regression (MLR): We use multi-layer regression with three fully connected layers and a hyperbolic tangent activation function as a deep learning-based regression.

Deep multi-layer perceptron (MLP): We use the multi-layer perceptron with four fully connected layers and a ReLU activation function as a deep learning-based multiclass classifier.

In the classification model that is Linear and MLP, we construct the classifier $f(\boldsymbol{x}) : \mathbb{R}^d \to \mathcal{L}$ where $\mathcal{L}$ is discretized location label set. In the regression model that is MLR, we construct the regressor $f(\boldsymbol{x}) : \mathbb{R}^d \to \mathcal{Y} \subset \mathbb{R}^2$ that estimates 2D coordinate directly. It should be noted that the CDFs of the MLR model are smoother than the other methods because MLR outputs the 2D coordinate directly; meanwhile, the output of the MLP and a linear model is discretized 2D location.

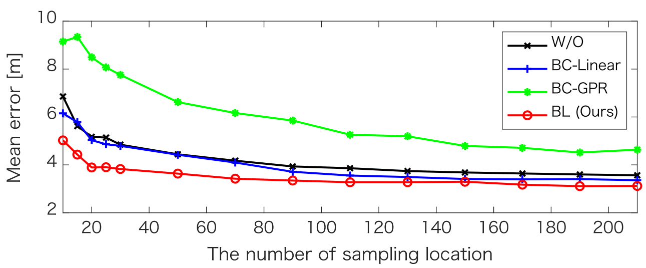
C. The number of data position vs performance

Fig. 7 shows the effect of changing the number of data sampling locations with mean localization errors. In this experiment, we randomly select the data sampling locations under 10 to 30 by increasing 5 locations and from 50 to 210 by increasing 20 locations. From this figure, our BL data augmentation method with 20 sampling locations achieves almost the same accuracy with the localization that using all sampling locations without data augmentation, in linear and MLP localization models, and with 70 sampling location for the MLR model. From this result, we confirmed our proposed method to improve the performance for the several localization models with 10% sampled location in the linear model and MLP model, and 30% sampled location in the MLR model.

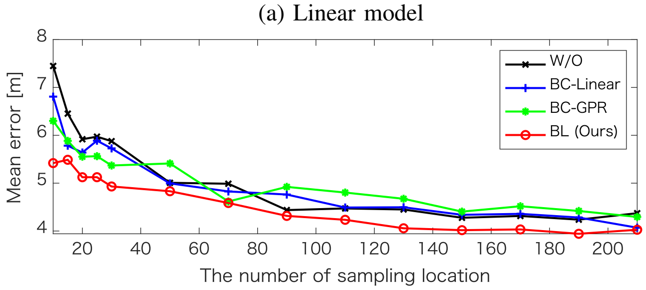
Moreover, our BL data augmentation most improves the localization performance than the other data augmentation methods at most settings of the number of sampling location and localization models. From the result, our data augmentation improved the performance on the settings from dense setting (210 sampled location) to extreme sparse setting (5 sampled locations). This result indicates that our BL data augmentation method generates the proper fingerprint data either the sampling location is sparse or not. In comparison with our method, linear and GPR data augmentation sometimes decreases the localization accuracy. This accuracy deterioration is because the improper fingerprint is generated from the GPR data augmentation.

TABLE I: Result on the errors in each setting

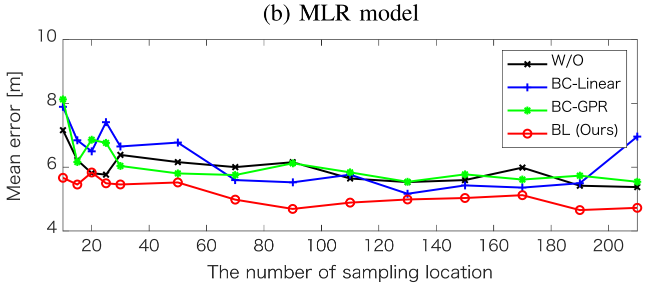
Localization	Data augmentation	Error on 210 sample locations [m]			Error on 83 sample locations [m]			Error on 30 sample locations [m]		
		Mean	90 percentile	Max	Mean	90 percentile	Max	Mean	90 percentile	Max
Linear	W/O	3.58	7.07	41.00	3.97	7.81	41.19	4.62	9.06	41.19
	BC-Linear	3.36	6.71	41.05	3.76	7.21	38.08	4.35	8.06	38.60
	BC-GPR	4.63	8.54	41.59	5.94	12.00	40.31	7.73	16.40	42.20
	BL (Ours)	**3.12**	**6.08**	**22.20**	**3.35**	**6.32**	**22.02**	**3.63**	**6.40**	**22.00**
MLR	W/O	4.37	7.49	24.10	4.46	7.86	24.32	6.92	12.73	34.45
	BC-Linear	4.07	7.35	22.13	4.62	8.47	27.27	5.14	9.17	**24.81**
	BC-GPR	4.29	7.63	21.25	4.93	8.97	24.52	5.31	9.55	30.29
	BL (Ours)	**4.03**	**7.04**	**20.08**	**4.28**	**7.80**	**22.08**	**4.29**	**7.93**	25.15
MLP	W/O	5.38	10.20	**38.05**	5.52	11.00	39.01	5.91	11.18	37.34
	BC-Linear	8.65	18.25	42.20	6.21	11.40	**37.01**	6.54	11.05	41.11
	BC-GPR	5.54	10.00	**38.05**	6.18	11.40	40.01	5.64	**10.05**	**31.00**
	BL (Ours)	**4.73**	**8.06**	40.00	**5.39**	**10.44**	39.12	**5.57**	**10.05**	39.00



(a) Linear model



(b) MLR model



(c) MLP model

Fig. 7: Localization error vs the number of location of training data

Additionally, the linear localization model outperforms the other localization method. This is because the linear model uses the loss function that evaluates the distance metric and employs the highly correlated feature representation to localization task, while the deep model should learn these properties from the limited dataset. It should be noted that this is out of the scope of our work, a more accurate model with deep learning would be constructed if we construct the indoor localization, specialized multi-layer model. When the accurate deep-learning model is realized, we can apply our BL data augmentation in the same manner.

D. Results of localization performance in several sampling location settings

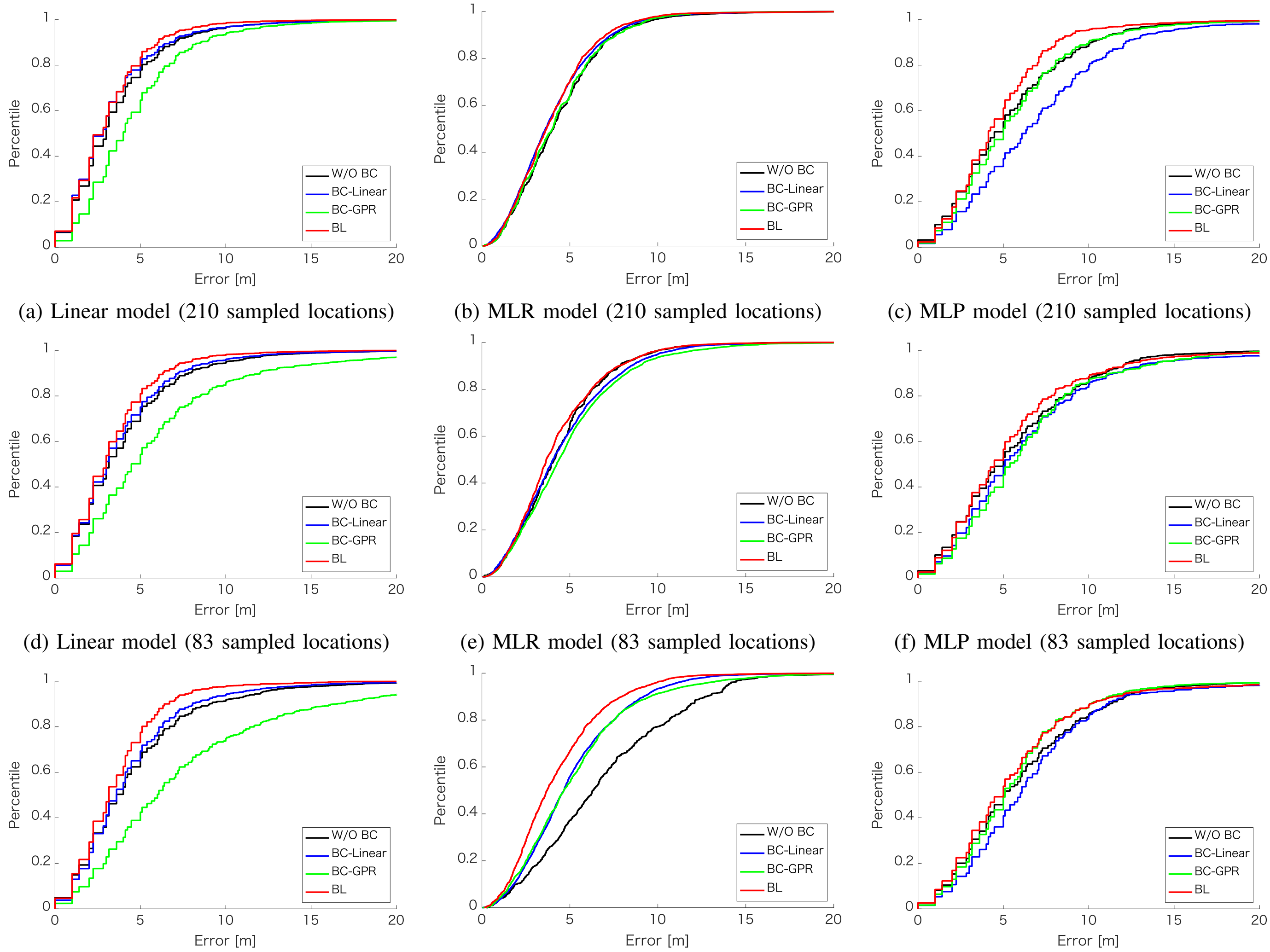
Here, we evaluate the localization performance for three settings: 1) sampling with all locations, 2) sampling at 83 locations with almost 2 m by 2 m, and 3) sampling at 30 locations with almost 3 m by 3 m. Table I is the error of mean, 90 percentile, and max error at each localization model and each data augmentation method on each setting. Fig. 8a to 8c show the CDFs of each data augmentation method in each localization model and each sampling setting. Moreover, we evaluate the RSSI radio map construction performance on our MLR model and GPR model in each setting.

1) Case 1: Dense setting with 210 sample location: Fig. 8a, 8b, and 8c show the CDF of the localization error with dense sampling setting in each localization method, respectively. From this result, our proposed model improves the 13 % in mean error, 14 % in 90 percentile, and 46 % at max error in linear model from the W/O result. In comparison with other localization methods, our BL data augmentation outperforms the other data augmentation method on accuracy improvement because the BL data augmentation, written in a red line, is placed on the most left upper side of the figure.

This result indicates that our proposed model improves the localization performance in the dense sampling setting. This is because generalization performance is improved owing to the increasing in a variety of data in each target location. Also, the result indicates that our BL data augmentation captures the RSSI distribution for the target environment precisely compared to other data augmentation methods when we use the dense dataset.

2) Case 2: Sparse setting with 83 sample locations: Fig. 8d, 8e, and 8f show the CDF of the localization error in sparse sampling setting with almost 2 m by 2 m. This result shows the localization performance in 1 m by 1 m that contains the 127 unseen location. From the figure, our method outperforms the other data augmentation method because our model is located on the most left upper side than other localization methods. From this result, our model improves the localization performance not only in dense settings but also sparsely sampled location settings.

Moreover, our method improves the mean error of the 6% in the linear model, 2% in the MLR model, and almost the same accuracy in the MLP model in case 2 setting from the



(a) Linear model (210 sampled locations)
(b) MLR model (210 sampled locations)
(c) MLP model (210 sampled locations)
(d) Linear model (83 sampled locations)
(e) MLR model (83 sampled locations)
(f) MLP model (83 sampled locations)
(g) Linear model (30 sampled locations)
(h) MLR model (30 sampled locations)
(i) MLP model (30 sampled locations)

Fig. 8: CDF on the errors in each setting

localization without data augmentation in dense setting (Case 1). Therefore, our model achieves similar accuracy to the localization model trained with the dense dataset with 40 % data sampling location.

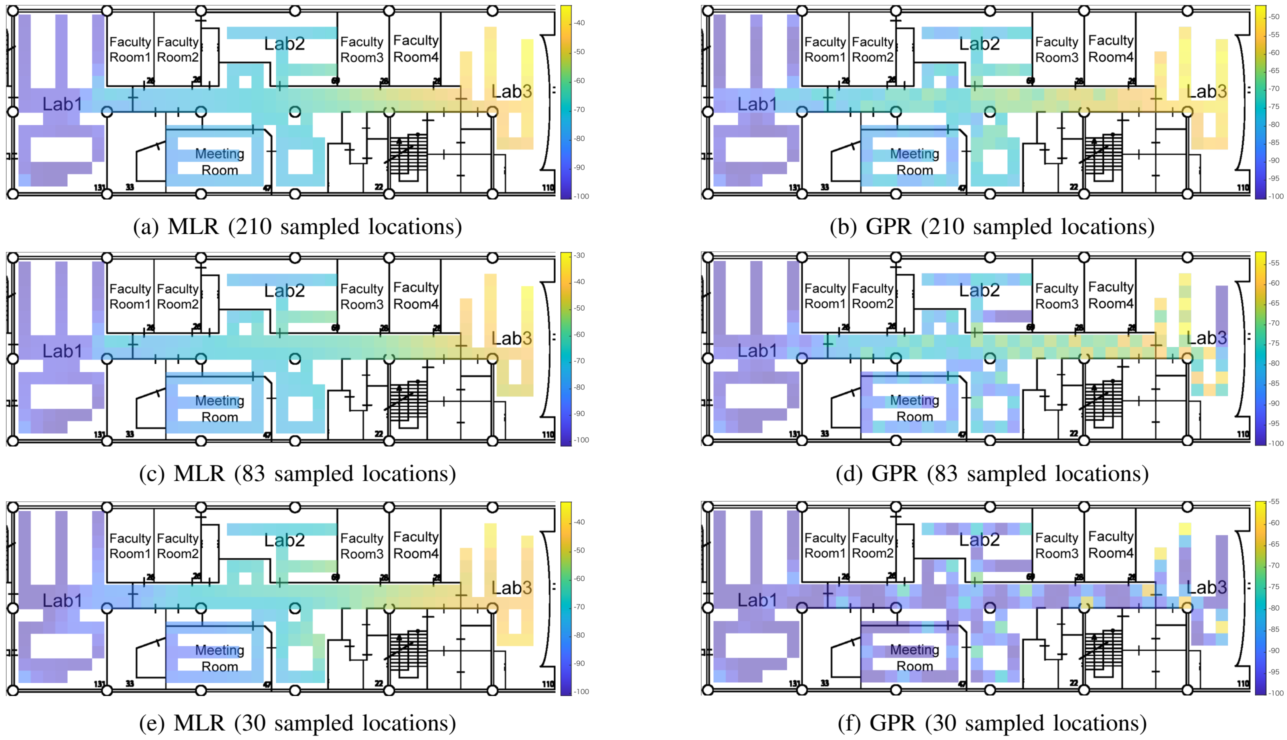
3) Case 3: Sparse setting with 30 sample locations: Fig. 8g, 8h, and 8i show the CDF of the localization error in sparse sampling setting with almost 3 m by 3 m. This result shows the localization performance in 1 m by 1 m that contains the 180 unseen location (only 15 % of the target location is sampled location). Even if this extreme sparse sampled location setting, our proposed model improve the localization performance at all localization method, while other data augmentation method less improves the performance or failure to improve the performance. This slight improvement or failure of improvement on comparison methods comes from the failure to generate the fingerprint data with actual fingerprint property. Our model generates the synthetic data that follows the actual fingerprint property with a sparse dataset compared to these models.

Compared with the localization without data augmentation in the dense setting, our method improves the -1% in the linear model, 2% in the MLR model, and -4% in MLP for the linear model in the sparse setting. From this result, slight accuracy degradation is observed in the linear and MLP model; however, we can regard almost the same accuracy with the dense setting in case 1. Therefore, our model achieves the accurate localization model only with sampling data from the 15 % of target locations.

TABLE II: Mean absolute error on RSSI in each setting

Data augmentation	210 sample	83 sample	30 sample
GPR	**8.71 [dBm]**	13.77 [dBm]	21.87 [dBm]
MLR (Ours)	**7.10 [dBm]**	**7.21 [dBm]**	**7.76 [dBm]**

4) Comparison of data generation in each setting: To evaluate the modeled RSSI radio map for data augmentation, we visualize the radio map modeled by our MLR model and GPR model. In this experiment, we selected the AP that has long propagation that set in the lab. 3. The ground truth is shown in Fig. 4a and AP location is shown in Fig. 10. Fig. 9a to Fig. 9f show the Wi-Fi RSSI radio map estimation result on MLP model and GPR model in each sampling setting at a specific access point. Table II shows the mean absolute error of each



(a) MLR (210 sampled locations)

(b) GPR (210 sampled locations)

(c) MLR (83 sampled locations)

(d) GPR (83 sampled locations)

(e) MLR (30 sampled locations)

(f) GPR (30 sampled locations)

Fig. 9: Modeled RSSI radio map

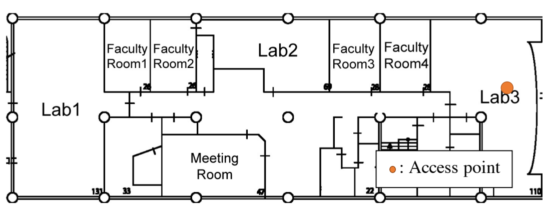


Fig. 10: Target access point location

method in each setting.

In the dense setting that is shown in Figs. 9a and 9b, the modeled radio maps in both model has similar trends and similar error value. From the fact that the RSSI has a 4 to 6 dBm noise discussed in the section III-C, both of the models can model the RSSI accurately. From the figures, our MLR model generated a more smooth RSSI radio map than the GPR model in a dense setting.

GPR model estimates the variance of the RSSI to generate the fingerprint in addition to the RSSI estimation and the mean absolute error of standard deviation was $2.27\ \mathrm{dBm}$. This result indicated that the GPR model has an error on the variances as well as RSSI estimation. This flexibility of variance estimation on GPR causes accumulation error at the data augmentation. Due to the accumulation error, the localization performance on GPR data augmentation got worse.

Fig. 9c depicts the generated radio map on our model in the 83 sampled location setting. From this figure, our model could generate a smooth RSSI radio map even if the sampling location is sparse. We can also observe that our model keeps the error on RSSI estimation at a similar level with the dense setting. Radio map generated by GPR that shown in 9d was affected sparse sampling location and RSSI estimation error also get worse and mean absolute error on the standard deviation is $2.78\ \mathrm{dBm}$. This difference comes from the model property of each method. MLR model is a parametric model that estimates RSSI radio map by using learned parameters from training data with the constraint of smoothness with a loss function. On the other hand, the GPR model is a non-parametric model that estimates RSSI radio maps by using training data directly. This means that the GPR model requires a dense sampling dataset for accurate radio map estimation and is unsuitable for sparse data modeling.

Figs. 9e and 9f also show the RSSI radio map estimation result on MLR and GPR model in 30 sampled location setting, respectively. The mean absolute error on the standard deviation for the GPR model is $2.72\ \mathrm{dBm}$. The result shows a similar trend with the 83 sampled location setting that our model estimated the smooth radio map with keeping the RSSI estimation error, and GPR got a worse result.

From the figure, our model successfully estimated the radio map accurately, although our model fails to estimate the peak of the RSSI distribution that is the AP location. This failure is caused because we did not set the sampling locations besides AP locations. However, from the localization errors discussed in the previous section, this failure has less effect on the localization error. From this observation, our model generated the RSSI radio map for improving the localization accuracy even if there is no sampling location besides AP locations.

V. Conclusion

We proposed a brand new data augmentation method for Wi-Fi fingerprint-based indoor localization named BL data augmentation, whatever the data collection location is dense or sparse. We first reviewed the BC learning and problems to apply that method to Wi-Fi fingerprint-based indoor localization. Then, we proposed the data generating methodology that consider the Wi-Fi signal property and label structure on indoor localization. Thanks to our method, we decrease the data sampling location and reduce the data acquisition cost, which is one obstacle for practical use of indoor localization.

In the experiment, we confirmed that BL data augmentation improves the localization performance in any indoor localization method. Also, the localization with BL data augmentation achieves the same accuracy by only using 10 % sampled location in the best case and 40 % in the worst case with localization without data augmentation using all sampled locations.

References

[1] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proc. of INFOCOM*, vol. 2, 2000, pp. 775–784.

[2] K. Ohara, T. Maekawa, and Y. Matsushita, "Detecting state changes of indoor everyday objects using Wi-Fi channel state information," *Proc. of IMWUT*, vol. 1, no. 3, 2017.

[3] C. Chen, Y. Chen, Y. Han, H.-Q. Lai, and K. R. Liu, "Achieving centimeter-accuracy indoor localization on WiFi platforms: A frequency hopping approach," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 111–121, 2016.

[4] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Proc. on DAR*, 2003, p. 958.

[5] J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. CVPR*, 2012, pp. 3642–3649.

[6] Y. Tokozume, Y. Ushiku, and T. Harada, "Between-class learning for image classification," in *Proc. on CVPR*, 2018, pp. 5486–5494.

[7] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization," *arXiv preprint arXiv:1710.09412*, 2017.

[8] R. Elbakly and M. Youssef, "A robust zero-calibration RF-based localization system for realistic environments," in *Proc. of SECON*, 2016.

[9] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using WiFi," in *Proc. of SIGCOMM*, vol. 45, no. 4, 2015.

[10] C. Zhang and X. Zhang, "Pulsar: Towards ubiquitous visible light localization," in *Proc. of MobiCom*, 2017.

[11] Z. Liu, L. Cheng, A. Liu, L. Zhang, X. He, and R. Zimmermann, "Multiview and multimodal pervasive indoor localization," in *Proc. of Multimedia*, 2017.

[12] S. P. Tarzia, P. A. Dinda, R. P. Dick, and G. Memik, "Indoor localization without infrastructure using the acoustic background spectrum," in *Proc. of MobiSys*, 2011.

[13] C.-Y. Hsu, A. Ahuja, S. Yue, R. Hristov, Z. Kabelac, and D. Katabi, "Zero-effort in-home sleep and insomnia monitoring using radio signals," *Proc. of IMWUT*, vol. 1, no. 3, p. 59, 2017.

[14] R. Kawajiri, M. Shimosaka, and H. Kashima, "Steered crowdsensing: incentive design towards quality-oriented place-centric crowdsensing," in *Proc. of UbiComp*, 2014.

[15] M. Shimosaka and O. Saisho, "Efficient calibration for RSSI-based indoor localization by Bayesian experimental design on multi-task classification," in *Proc. on Ubicomp*, 2016, pp. 244–249.

[16] C. Gleason, D. Ahmetovic, S. Savage, C. Toxtli, C. Posthuma, C. Asakawa, K. M. Kitani, and J. P. Bigham, "Crowdsourcing the installation and maintenance of indoor localization infrastructure to support blind navigation," *Proc. of IMWUT*, vol. 2, no. 1, p. 9, 2018.

[17] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. R. Choudhury, "No need to war-drive: unsupervised indoor localization," in *Proc. of Mobisys*, 2012.

[18] Y. Jiang, X. Pan, K. Li, Q. Lv, R. P. Dick, M. Hannigan, and L. Shang, "ARIEL: Automatic Wi-Fi based room fingerprinting for indoor localization," in *Proc. of UbiComp*, 2012.

[19] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proc. of MobiCom*, 2012.

[20] Y. Jiang, Y. Xiang, X. Pan, K. Li, Q. Lv, R. P. Dick, L. Shang, and M. Hannigan, "Hallway based automatic indoor floorplan construction using room fingerprints," in *Proc. of Ubicomp*, 2013.

[21] H. Kashima *et al.*, "A semi-supervised approach to indoor localization estimation," in *IEEE ICDM Data Mining Contest 2007*.

[22] B. Chidlovskii and L. Antsfeld, "Semi-supervised variational autoencoder for WiFi indoor localization," in *Proc. of IPIN2019*.

[23] M. Zhou, Y. Tang, W. Nie, L. Xie, and X. Yang, "GrassMA: graph-based semi-supervised manifold alignment for indoor WLAN localization," *Trans. on IEEE Sensors*, vol. 17, no. 21, pp. 7086–7095, 2017.

[24] J. Bi, Y. Wang, Z. Li, S. Xu, J. Zhou, M. Sun, and M. Si, "Fast radio map construction by using adaptive path loss model interpolation in large-scale building," *Trans. on MDPI Sensors*, vol. 19, no. 3, p. 712, 2019.

[25] M. K. Hoang, J. Schmalenstroeer, and R. Haeb-Umbach, "Aligning training modelswith smartphone properties in WiFi fingerprinting based indoor localization," in *Proc. of ICASSP2015*, pp. 1981–1985.

[26] X. Wang, X. Wang, S. Mao, J. Zhang, S. C. Periaswamy, and J. Patton, "DeepMap: Deep Gaussian process for indoor radio map construction and location estimation," in *Proc. of GLOBECOM2018*.

[27] H. Zou, M. Jin, H. Jiang, L. Xie, and C. J. Spanos, "WinIPS: WiFi-based non-intrusive indoor positioning system with online radio map construction and adaptation," *Tran. on Wireless Communications*, vol. 16, no. 12, pp. 8118–8130, 2017.

[28] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *Trans. on IEEE Access*, vol. 3, pp. 1089–1100, 2015.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. of NeurIPS*, 2012, pp. 1097–1105.

[30] A. Dosovitskiy, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative unsupervised feature learning with convolutional neural networks," in *Proc. of NeurIPS*, 2014, pp. 766–774.

[31] B. Ferris, D. Fox, and N. D. Lawrence, "WiFi-slam using Gaussian process latent variable models." in *Proc. of IJCAI*, vol. 7, no. 1, 2007, pp. 2480–2485.

[32] H.-C. Yen and C.-C. Wang, "Adapting Gaussian processes for cross-device Wi-Fi localization," in *Proc. of IPIN*, 2013, pp. 1–8.

[33] S. Sen, J. Lee, K.-H. Kim, and P. Congdon, "Avoiding multipath to revive inbuilding WiFi localization," in *Proc. of MobiSys*, 2013, pp. 249–262.

[34] M. Sugasaki and M. Shimosaka, "Robust indoor localization across smartphone models with ellipsoid features from multiple RSSIs," *Proc. of IMWUT*, vol. 1, no. 3, p. 103, 2017.

Masato Sugasaki received his BE and ME from the Tokyo Institute of Technology in 2016 and 2018, respectively. He is currently a Ph.D. candidate student at the Tokyo Institute of Technology. His research interests include indoor localization, ubiquitous computing.

Masamichi Shimosaka received his BE, ME, and Ph.D. from the University of Tokyo in 2001, 2003, and 2006, respectively. He joined Tokyo Institute of Technology as an associate professor in July 2015. Prior to joining Tokyo Institute of Technology, he was a faculty member at the University of Tokyo from 2006 to 2015. While his Ph.D. candidate, he was funded by the Japan Society for the Promotion Science as a research fellow. His research interests include machine intelligence and ubiquitous computing. He is a member of the ACM and IEEE.