Deep Learning-based Localization Using Spatial Information in the NLoS Scenarios

Yongjun Ahn, Inkook Keum, Jinwoo Son, and Byonghyo Shim
Institute of New Media and Communications and Department of Electrical and Computer Engineering,
Seoul National University, Korea
Email: {yjahn, ikkeum, jinwooson, bshim}@islab.snu.ac.kr

***Abstract*—In this paper, we propose a new type of deep learning (DL)-based localization for the urban NLoS scenarios, termed *Intelligent Localization via Spatial Information Embedding* (I-LOSIE). In I-LOSIE, we train the deep neural network (DNN) to identify the location of the wireless device by using the combination of the uplink measurements and the position and size of the obstacles (we call these *spatial information*). From the extensive numerical evaluations, we show that the proposed I-LOSIE achieves the high-resolution localization in the wireless environments with high density of obstacles..**

I. Introduction

With the advent of the internet of things (IoT) era, location awareness, providing the ability to identify the location of sensor, machine, vehicle, and wearable device, has become one of key ingredients for the hyper-connected society [1]. This new trend has promoted the emergence of new killer applications, such as autonomous driving, smart factory/monitoring, drone delivery, and remote surgery, that require an extremely precise localization capability down to several centimeters for the operation [2]–[4].

Traditionally, trilateration-based techniques such as the global navigation satellite systems (GNSS) have been widely used for the user localization [5]–[7]. In these approaches, the distances between the user and sensor nodes are obtained from the signal measurements such as time-of-arrival (ToA) and received signal strength indicator (RSSI). By identifying the intersection point of the spheres centered at the sensor nodes, the position of the target device is estimated. While this approach is easy and straightforward, it does not perform well when the wireless signal is propagated through the non-line-of-sight (NLoS) paths. For instance, when the transmit signal is reflected by a scatterer in the middle of propagation, the measured distance becomes longer than that of the LoS path (i.e., actual distance). In this case, the volume of the sphere becomes unduly large. This increases the size of intersection area in the trilateration-based techniques, resulting in a high positioning error.

An aim of this paper is to propose a novel localization scheme for the NLoS propagation scenarios using deep learning (DL). We employ DL to extract the geometric relation between the obstacles/scatterers and the blockages/reflections

This work was supported by the National Research Foundation of Korea (NRF) Grant through the Ministry of Science and ICT (MSIT), Korea Government, under Grant 2022R1A5A1027646.

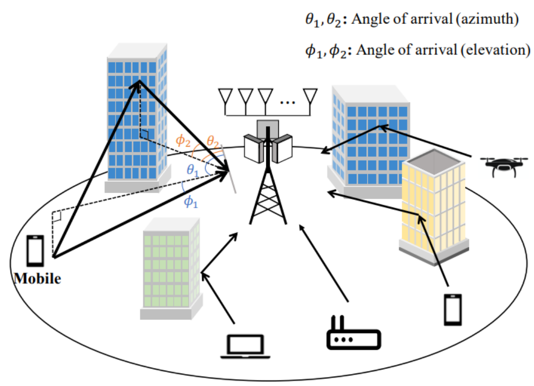

Fig. 1. An illustration for narrowband uplink transmission scenario.

determining the type of signal propagation (i.e., LoS or NLoS). Specifically, in our framework, henceforth referred to as *Intelligent Localization via Spatial Information Embedding* (I-LOSIE), the deep neural network (DNN) is trained to find out the location of the user based on two inputs; 1) the position and size of the obstacles (i.e., spatial information) and 2) the uplink signal measurements (i.e., channel information). By extracting the propagation characteristics (e.g., directions of signal reflection, areas where the blockage occurs), I-LOSIE can identify the type of signal propagation (i.e., LoS or NLoS) and the position where the user is most likely to exist, achieving a high-resolution localization.

II. Uplink System Model for Localization

We consider a narrowband uplink transmission scenario where a base station (BS) equipped with $N_T = N_x \times N_y$ uniform planar array (UPA) antennas serves a single-antenna mobile (see Fig. 1). In this setup, the received pilot signal $\mathbf{y} \in \mathbb{C}^{N_T \times 1}$ is given by

$$\mathbf{y} = \mathbf{h}s + \mathbf{n}, \tag{1}$$

where $\mathbf{h} \in \mathbb{C}^{N_T \times 1}$ is the uplink channel vector from the mobile to the BS, $s \in \mathbb{C}$ is a transmitted pilot symbol, and $\mathbf{n} \in \mathbb{C}^{N_T \times 1}$ is the additive Gaussian noise ($\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$).

As for the channel model, we consider a geometric channel model where the uplink channel vector $\mathbf{h}$ is expressed as

$$\mathbf{h} = \sum_{i=1}^{N_p} \alpha_i e^{-j2\pi f_s k \tau_i} \mathbf{a}(\theta_i, \phi_i), \tag{2}$$

where N_p is the number of propagation paths, α_i is complex gain of the i-th path, θ_i and ϕ_i are azimuth and elevation angles for the i-th path, and τ_i is the path delay of the i-th path. Also, $\mathbf{a}(\theta,\phi) \in \mathbb{C}^{N_T \times 1}$ is the array response of BS given by

$$\mathbf{a}(\theta,\phi) = \frac{1}{\sqrt{N_T}}[1 \cdots e^{-j\pi(N_x-1)\cos\theta\sin\phi}]^{\mathrm{T}} \otimes [1 \cdots e^{-j\pi(N_y-1)\sin\theta\sin\phi}]^{\mathrm{T}}. \quad (3)$$

We assume that the AoAs $\{\theta,\phi\}$ and distance $d = c\tau$ of the propagation path can be accurately obtained from $\mathbf{y}$, where c is the speed of light[1]. Based on the assumption, we solve the localization problem that finds out the position of the mobile $\mathbf{p}_m = [x,y,z]$ from the AoAs $\{\theta,\phi\}$ and distance d.

III. Deep Learning-based 3D Localization

The primary goal of I-LOSIE is to estimate the 3D location of a mobile from the channel information and spatial information. Major benefit of I-LOSIE is to exploit the position and size of the obstacles (we call these *spatial information*). By extracting the signal propagation geometry (e.g., directions of signal reflection, areas where the blockage occurs) from the spatial information, I-LOSIE can distinguish the type of signal propagation (i.e., LoS or NLoS) and also identify the 3D position of the user accurately. Specifically, we exploit the fully-connected (FC) network consisting of multiple hidden layers to learn the input-output mapping g [10]:

$$\hat{\mathbf{p}}_m = g(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, \mathbf{s}_1, \cdots, \mathbf{s}_K; \delta), \quad (4)$$

where $\mathbf{r}_n = [d_n, \theta_n, \phi_n]$ is the set of distance and angles of n-th path (i.e., channel information) and δ is set of weights and biases. Also, $\mathbf{s}_k = [x_k, y_k, w_k, l_k, h_k]$ is the set of the center coordinates (x_k, y_k) and width w_k, length l_k, and height h_k of the k-th obstacle (i.e., spatial information).

A. I-LOSIE Architecture

Main goal of I-LOSIE is to find out the exact location of the user from the spatial information $\mathbf{s}$ and channel information $\mathbf{r}$. To this end, we train a DNN consisting of multiple FC layers. Specifically, in I-LOSIE, we first construct the vector $\mathbf{x}^{(0)} = [\mathbf{r}_1^{\mathrm{T}}, \cdots, \mathbf{r}_N^{\mathrm{T}}, \mathbf{s}_1^{\mathrm{T}}, \cdots, \mathbf{s}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{(3N+5K)\times 1}$ by concatenating the N multipath inputs and K spatial inputs. By using $\mathbf{x}^{(0)}$ as the input, the first FC layer generates the output $\mathbf{z}^{(0)} \in \mathbb{R}^{E\times 1}$ as

$$\mathbf{z}^{(0)} = \mathbf{W}^{(0)}\mathbf{x}^{(0)} + \mathbf{b}^{(0)}, \quad (5)$$

where $\mathbf{W}^{(0)} \in \mathbb{R}^{E\times(3N+5K)}$ and $\mathbf{b}^{(0)} \in \mathbb{R}^{E\times 1}$ are the weight matrix and the bias vector, respectively. After passing through the FC layer, the batch normalization is applied to alleviate the large variation of the inputs caused by the different channel state and noise level [11]. To be specific, the j-th element of mini-batch $\mathbf{B} = [\mathbf{z}^{(0),1}, \cdots, \mathbf{z}^{(0),b}]$ after the batch normalization is

$$\tilde{z}_i^{(0),j} = \gamma\left(\frac{z_i^{(0),j} - \mu_{\mathbf{B},i}}{\sqrt{\sigma_{\mathbf{B},i}^2}}\right) + \beta, \quad i = 1, \cdots, E, \quad (6)$$

where $\mu_{\mathbf{B},i} = \frac{1}{b}\sum_{j=1}^{b} x_i^{(0),j}$ and $\sigma_{\mathbf{B},i}^2 = \frac{1}{b}\sum_{j=1}^{b}(x_i^{(0),j} - \mu_{\mathbf{B},i})^2$ are mini-batch-wise mean and variance, respectively, γ is the scaling parameter, and β is the shifting parameter. After the batch normalization process, the output vector $\check{\mathbf{z}}^{(0)} = f_{\mathrm{ReLU}}(\tilde{\mathbf{z}}^{(0)})$ is generated by passing through the rectified linear unit (ReLU) layer $f_{\mathrm{ReLU}} = \max(0,x)$.

Then, the output vector $\check{\mathbf{z}}^{(0)}$ passes through L series of FC layers, batch normalization layers, and activation layers, generating the output vector $\check{\mathbf{z}}^{L}$. Finally, using $\check{\mathbf{z}}^{L}$, we obtain the location estimate of the mobile user $\hat{\mathbf{p}}_m = [\hat{x},\hat{y},\hat{z}]$:

$$[\hat{x},\hat{y},\hat{z}] = \mathbf{W}^f\check{\mathbf{z}}^L + \mathbf{b}^f, \quad (7)$$

where $\mathbf{W}^f$ and $\mathbf{b}^f$ are weight and bias.

B. Loss Function Design and I-LOSIE Training

In the training phase, we update the network parameters in the direction of minimizing the loss function $\mathcal{L}_\delta$. To obtain the optimal mapping g^* from the I-LOSIE training, we need to compare the 3D location estimate $\hat{\mathbf{p}}_m$ against the true location $\mathbf{p}_m$. The loss function $\mathcal{L}_\delta$ quantifying the difference between the true and estimated locations is given by the mean squared error (MSE):

$$\mathcal{L}_\delta = \|\hat{\mathbf{p}}_m - \mathbf{p}_m\|_2^2. \quad (8)$$

It is worth mentioning that the proposed I-LOSIE is the data-driven approach. This means that the DNN should be re-trained in the new wireless environments where the distribution of the input-output pairs is dearly distinct from the original dataset. For example, when the communication environment changes from the indoor to outdoor, the average communication distance can vary (e.g., 4 m in indoor and 9 m in outdoor) so that I-LOSIE trained in the indoor environment might not properly capture the wireless geometry of the outdoor environment [12]. To address the issue, we exploit the *meta learning*, a technique to train a model on various tasks such that it can solve new task using only a small number of training samples. In short, meta learning is a special training technique to obtain the initialization parameters of DNN using which one can easily and quickly learn the desired function (in our case, the mapping function g) with a few training samples.

Overall procedure of the I-LOSIE training is as follows. First, we perform the meta learning to obtain the initialization parameters. We then update the network parameters to perform the fine-tuning of DNN such that the trained DNN approximates the mapping g for the desired wireless environments (see Fig. 2). To be specific, the network parameter

[1]For example, we can accurately extract $\{\theta,\phi\}$ and d from $\mathbf{y}$ by using the CS-based technique that converts the angle and path delay estimation problem to the support identification problem [8], [9].

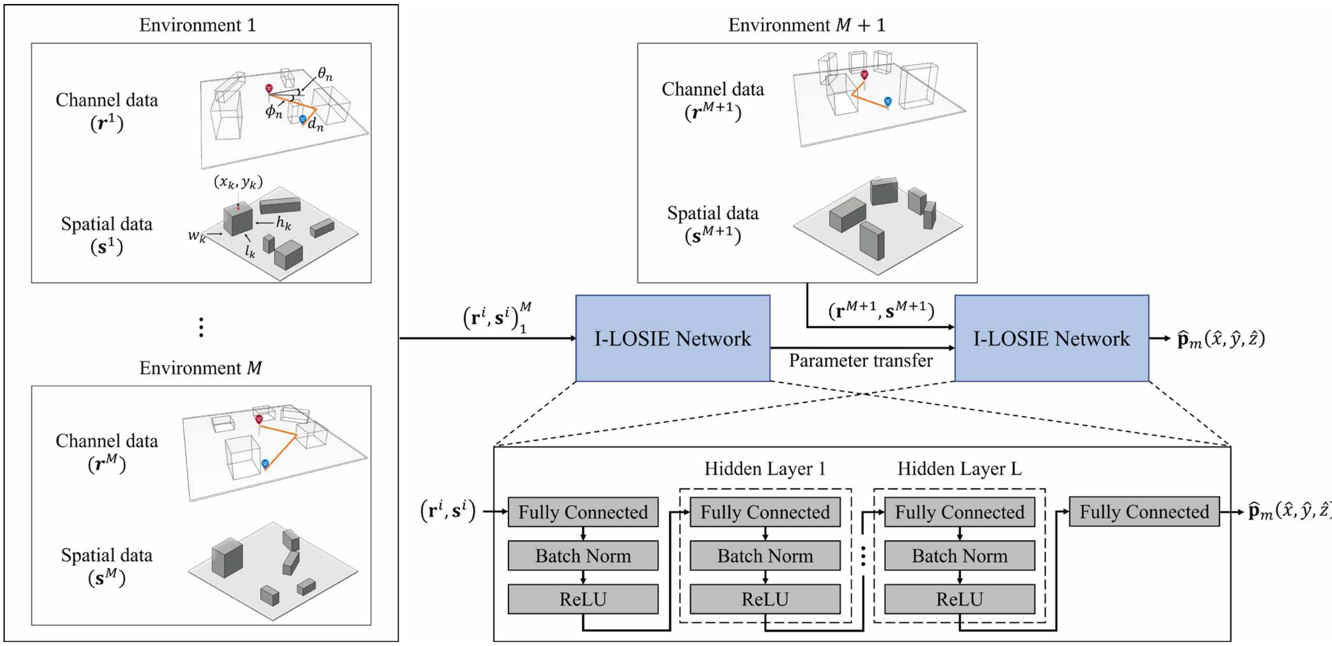

Fig. 2. Training process of I-LOSIE including the detailed structure.

update process in the meta learning phase using M datasets $\{D_1, \cdots, D_M\}$ is [13]

$$\psi_{D_i,t} = \delta_{t-1} - \alpha \nabla_\delta \mathcal{L}_{\delta_{t-1}}^{D_i}, \tag{9}$$

$$\delta_t = \delta_{t-1} - \beta \nabla_\delta \sum_{i=1}^{M} \mathcal{L}_{\psi_{D_i,t}}^{D_i}, \tag{10}$$

where δ_t is the set of network parameters updated at t-th step and $\psi_{D_i,t}$ is the set of network parameters temporally computed with dataset D_i at t-th step. Also, $\mathcal{L}^{D_i}$ is the loss function of the DNN for i-th dataset D_i. In the fine-tuning phase, we utilize δ as the initial parameter of I-LOSIE. Since all we need in the fine-tuning is to learn the distinct features of D_{M+1} unextracted from the meta learning, we can greatly reduce the training overhead.

IV. Simulations and Discussions

A. Simulation Setup

In our simulations, we consider the rectangular outdoor environment of $140\,\text{m} \times 140\,\text{m}$. We generate 4 different datasets $\{D_i\}_{i=1}^4$ for meta learning, each of which corresponds to different obstacle displacement. For all environments, we set the height of the BS to $10\,\text{m}$ and locate the BS at the center of the site. User devices are uniformly distributed in the service area. For the I-LOSIE structure, we employ the FC network consisting of 6 FC layers. Each hidden layer consists of $1,024$ hidden units. In the meta learning phase, we train I-LOSIE for $110,000$ iterations using $10,000$ samples.

B. Experiment Results and Discussions

In Table I, we show the average positioning error of the I-LOSIE with and without the spatial information. In this simulation, we use $1,000$ samples for the fine-tuning in the new environment. We observe that the proposed I-LOSIE achieves the localization performance gain when the spatial information is given as the additional input. For example, we see that the average localization error is decreased by about 43% for the single path and 55% for the multipath, respectively. This is mainly because the model can resolve the ambiguity of the LoS/NLoS propagation by utilizing the additional blockage and reflection information given from the spatial input.

TABLE I
Localization error in terms of MAE for the single path and multipath scenarios with and without spatial information $\mathbf{s}$.

Average (m)	single path	multipath
Without spatial info.	1.421	1.198
With spatial info.	0.811	0.536

Fig. 3 depicts MAE performance of the I-LOSIE trained with and without meta learning process as a function of the number of samples used for fine tuning. As shown in Fig. 3, MAE of the I-LOSIE without meta learning is significantly higher than that of I-LOSIE with meta learning. For example, when the number of samples is $6,700$, the MAE of I-LOSIE without and with meta learning is $1.95\,\text{m}$ and $0.54\,\text{m}$, respectively. This is mainly because the I-LOSIE with meta learning is pre-trained with sufficient amount of data to learn the essential knowledge needed for the NLoS localization (e.g., use of spatial information in the NLoS scenario in calculating the signal reflection).

V. Conclusion

In this paper, we proposed a novel DL-aided localization technique called I-LOSIE for the NLoS propagation scenarios.

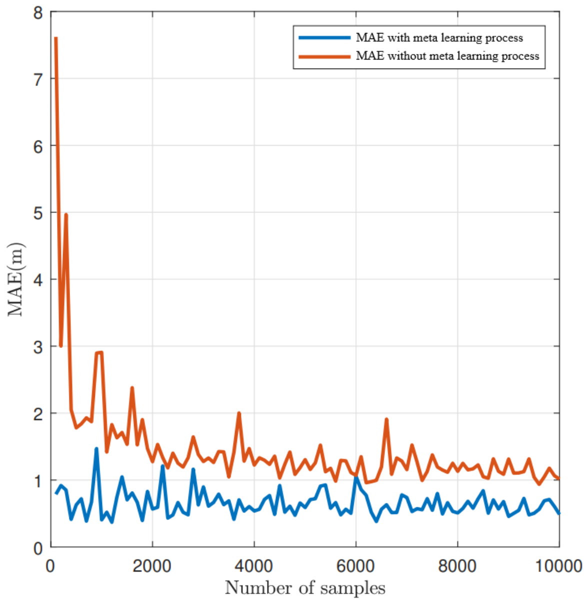

Fig. 3. MAE with respect to the number of fine-tuning samples.

Key idea of the proposed I-LOSIE is to identify the type of signal propagation by extracting the geometric information (e.g., direction of signal reflection) from the spatial data describing the size and position of obstacles. By training DNN via meta learning, we obtain I-LOSIE that can achieve an accurate localization using only a small number of training samples. From the extensive simulations in the various NLoS propagation scenarios, we confirmed that the proposed I-LOSIE is applicable for the high-resolution localization in the wireless environments with high density of obstacles.

REFERENCES

[1] L. T. Nguyen, J. Kim, S. Kim, and B. Shim, “Localization of iot networks via low-rank matrix completion,” *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5833–5847, 2019.

[2] F. Zafari, A. Gkelias, and K. K. Leung, “A survey of indoor localization systems and technologies,” *IEEE Commun. Surv. Tutor.*, vol. 21, no. 3, pp. 2568–2599, 2019.

[3] I. Guvenc and C.-C. Chong, “A survey on toa based wireless localization and nlos mitigation techniques,” *IEEE Commun. Surv. Tutor.*, vol. 11, no. 3, pp. 107–124, 2009.

[4] Y. Ahn, J. Kim, S. Kim, K. Shim, J. Kim, S. Kim, and B. Shim, “Towards intelligent millimeter and terahertz communication for 6g: Computer vision-aided beamforming,” *IEEE Wireless Commun.*, 2022.

[5] L. T. Nguyen, J. Kim, S. Kim, and B. Shim, “Localization of IoT networks via low-rank matrix completion,” *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5833–5847, 2019.

[6] H. C. So, Y. T. Chan, and F. K. W. Chan, “Closed-form formulae for time-difference-of-arrival estimation,” *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2614–2620, 2008.

[7] A. Zanella, “Best practice in RSS measurements and ranging,” *IEEE Commun. Surv. Tutor.*, vol. 18, no. 4, pp. 2662–2686, 2016.

[8] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, “Compressed sensing for wireless communications: Useful tips and tricks,” *IEEE Commun. Surv. Tutor.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[9] Y. Ahn, W. Kim, and B. Shim, “Active user detection and channel estimation for massive machine-type communication: Deep learning approach,” *IEEE Internet Things J.*, vol. 9, no. 14, pp. 11 904–11 917, 2021.

[10] S. Kim, J. Son, and B. Shim, “Energy-efficient ultra-dense network using lstm-based deep neural networks,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4702–4715, 2021.

[11] W. Kim, Y. Ahn, and B. Shim, “Deep neural network-based active user detection for grant-free noma systems,” *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2143–2155, 2020.

[12] J. Kim, Y. Ahn, and B. Shim, “Massive data generation for deep learning-aided wireless systems using meta learning and generative adversarial network,” *IEEE Trans. Vehi. Technol.*, vol. 72, no. 1, pp. 1302–1306, 2022.

[13] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” in *Proc. Int. Conf. Mach. Learn. (ICML)*. PMLR, 2017, pp. 1126–1135.