Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9680479, 10 pages
https://doi.org/10.1155/2022/9680479

Research Article

Deep Learning-Based Localization with Urban Electromagnetic and Geographic Information

Wenyu Wang,[1] **Baozhu Li**,[2] **Zhen Huang**,[2] **and Lei Zhu**[1]

[1]*College of Communications Engineering, Army Engineering University of PLA, Nanjing 210001, China*
[2]*Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100089, China*

Correspondence should be addressed to Zhen Huang; huangzhen@tsinghua.edu.cn and Lei Zhu; zhulei_paper@126.com

Received 12 May 2022; Revised 1 August 2022; Accepted 11 August 2022; Published 21 August 2022

Academic Editor: Mingqian Liu

There is a growing demand for localization of illegal signal sources, aiming to guarantee the security of urban electromagnetic environment. The performance of traditional localization methods is limited due to the non-line-of-sight (NLOS) propagation and sparse layouts of sensors. In this paper, a deep learning-based localization method is proposed to overcome these issues in urban scenarios. Firstly, a model of electromagnetic wave propagation considered with geographic information is proposed to prepare reliable datasets for intelligent cognition of urban electromagnetic environment. Then, this paper improves an hourglass neural network which consists of downsampling and upsampling layers to learn the propagation features from sensing data. The core modules of VGG and ResNet are, respectively, utilized as feature extractors in downsampling. Moreover, this paper proposes a weighted loss function to expand the attention on position features, in order to improve the performance of localization with sparse layouts of sensors. Representative numerical results are discussed to assess the proposed method. ResNet-based extractor performs more efficiently than VGG-based extractor, and the proposed weighted loss function increases the localization accuracy by more than 50%. Additionally, the established geographic model supports qualitative and quantitative evaluation of the robustness with varied degree of NLOS propagation. Compared with other deep learning-based algorithms, the proposed method presents the more robust and superior performance under severe NLOS propagation and sparse sensing conditions.

1. Introduction

Along with the acceleration of social urbanization and informatization progresses, human activities have been concentrated in urban areas. The urban electromagnetic environment has become much more complex as well. Illegal signal sources in the cities encroach the spectrum resources without a valid license or even cause harmful interference, which probably leads to the loss of data and other critical faults in the communication systems [1]. As a result, localization of illegal signal sources is significant to guarantee the security of the urban electromagnetic environment. Wireless sensor networks (WSN) [2] utilize a set of sensors to monitor the electromagnetic activities in the target area. These sensors are characterized by small volume and low power, working in concert for localization by collecting and processing the signal parameters such as time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), and received signal strength (RSS) [3].

However, it is unavoidable for the reflections and diffractions of electromagnetic wave to create non-line-of-sight (NLOS) propagation due to the dense buildings in urban scenarios. Compared with line-of-sight (LOS) propagation, it results in longer propagation distance and time, so that the localization errors reach to hundreds of meters [4]. In this case, traditional geometry-based positioning methods require at least three LOS paths between sensors and signal sources to solve nonlinear equations (such as TOA-based, TDOA-based, AOA-based, and RSS-based equations) which are susceptible to severe NLOS propagation conditions [5–8].

Addressing the challenges of localization in urban environment, researchers focus on the reduction of the estimation errors for signal parameters by identifying and discarding the sensing data of NLOS or changing the weight of NLOS data

[9–13]. These solutions, in essence, still apply LOS paths for localization as much as possible, but LOS paths are not available in complex urban scenarios. Furthermore, there are also some approaches (e.g., [14–16]) mitigating the NLOS errors by nonlinear optimization, convex relaxation, and other algorithms with prior information. A solution based on robust least squares and semidefinite relaxation is suggested in [16]. It is able to carry out the localization with only one LOS path but requires prior knowledge of the upper bound loss in NLOS propagation. However, the prior knowledge is difficult to obtain in the complex electromagnetic environment. Additionally, with the development of modern geographic information system (GIS) techniques, geographic models are introduced into relevant field [17, 18]. A floor map [19] is used to initialize and update the landmark graph for indoor localization with the aid of multifunction sensing components. In [20], the environmental information is imported into the proposed quasi 3D ray-tracing model to construct the synthetic radio map, and then, the refined radio map is used to prepare the training datasets for the indoor intelligent target intrusion sensing and localization. Focusing on the localization under NLOS conditions, Zhang et al. [21] correct the NLOS errors with the aid of 3D mapping database and employ factor graph optimization for the positioning calculation in global navigation satellite system. Perez-Cruz et al. [22] propose a probabilistic algorithm to learn and correct the NLOS biases. The above approaches are all generalized by two steps, i.e., calibrate the parameter estimation errors caused by NLOS propagation and solve the positioning equations by various optimization algorithms. Undeniably, such solutions are mathematically intractable and faced with double challenges from parameter estimation errors and localization errors.

From another perspective, since the NLOS propagation channels are relatively stable in a certain urban area with known geographic information, solutions of localization can be formulated to seek for the mapping of signal positions with sensing data by learning the NLOS propagation features in the target area, which offers an end-to-end intelligent cognition service. Deep learning has been proved to be an efficient tool in a wide range of fields due to its outstanding capability to capture the features and learn the mapping of data [23, 24]. Recently, several researches [25–27] apply deep neural networks (DNN) for end-to-end localization and gain some improvement. A long short-term memory (LSTM) network is used in [25] for small-scale indoor localization. Lin et al. [26] suggest a heatmap regression-based HMTLNet, revealing the active role of ResNet in a fully convolutional network (FCN) for localization. Zhan et al. [27] present a convolutional neural network (CNN) to solve localization problems from a view of computer vision. However, there are mainly two challenges of these methods.

(1) The empirical and statistical propagation model used in simulated validation (e.g., [25, 26]) is far from the actual electromagnetic propagation data, without considering the geographical distribution in urban areas

(2) The existing deep learning-based localization methods require massive data and dense deployment of sensors for perception of the target area. For example, the density of sensors in DeepMTL method [27] is 2 percent, with the number of sensors reaching to 200 in a $1\,\text{km} \times 1\,\text{km}$-sized area, which is costly in practice

To adapt to the NLOS propagation and sparse layouts of sensors, a localization method based on the intelligent cognition of urban electromagnetic environment is proposed in this paper. The main contribution is that geographic information is considered in the electromagnetic wave propagation model. The model prepares datasets which are closer to actual electromagnetic propagation. More specifically, an hourglass neural network and a weighted loss function are proposed to predict the position probability distribution with the sparse sensing data in urban electromagnetic environment. Finally, the performance of the proposed method is verified on the reliable datasets with extremely sparse sensors.

The rest of this paper is organized as follows. In Section 2, reliable datasets are prepared with the aid of geographic information. In Section 3, the proposed deep learning-based localization method is depicted in detail. In Section 4, the comparative simulations with different density of sensors are presented and discussed. Finally, conclusions are stated in Section 5.

2. Data Preparation

In this section, we introduce the urban propagation model considered with geographic information and the prepared datasets for intelligent cognition.

2.1. The Urban Propagation Model with Geographic Information. NLOS propagation is mainly caused by ray blocking from varied objects in urban scenarios. The layouts and positions of urban objects are essential on the analysis of NLOS propagation. Generally, the GIS data are open access, which contains the distribution of buildings, vegetation, rivers, and other objects. Based on the GIS information, the geographic model G is established. This paper is aimed at learning features in NLOS propagation with the aid of geographic information, demanding of massive data. Calculating accurate electromagnetic wave propagation is the core step to prepare effective datasets for deep learning-based localization. Empirical or statistical electromagnetic propagation models (e.g., COST 231 Walfisch-Ikegami propagation model [28]) only consider statistical features of different geographic scenarios roughly but take no account of the specific geographic models of the study area. Hence, they are prone to deviate a lot from practical propagation. According to the interactions between electromagnetic waves and urban geographical objects, ray-tracing model [29] is able to track hundreds of propagation paths for each receiving sensor but is susceptible to small errors in geographic models. Considering geographic information, Wahl et al. [30] propose dominant path model (DPM). In [31], DPM has been proved as a more accurate and robust electromagnetic propagation model than intelligent ray tracing

(IRT) model and the COST 231 Walfisch-Ikegami model in complex urban scenarios. Therefore, in this paper, DPM is utilized to calculate the propagation loss based on the established geographic model G.

Assume that the illegal signal source is located at $s_i = (x_i, y_i)$, and coordinates of all receiving points (different from the definition of sensor positions) in G are collected in the set $M = \{m_1, m_2, \cdots, m_Q\}$. Then, the length of the propagation path between the signal source and a receiving point is formulated as $d(s_i, m_j \in M(s_i))$, which also includes the case that the direction of propagation is changed by blocking. The function $f(\varphi, k)$ in dB means the kth interaction loss with the new propagation direction φ. Therefore, the predicted loss in G is calculated as

$$L(s_i, m_j) = -27.56 + 20\log(f) + 10p\log(d(s_i, m_j)) + \sum_{k=0}^{n} f(\varphi, k) - \frac{1}{c}\sum_{u=0}^{c} w_u. \quad (1)$$

Here, f is the electromagnetic wave frequency in MHz. The factor p depends on the LOS or NLOS condition of the propagation path. And w is the waveguiding factor which represents the effect of reflection along the walls in dominant path.

Since RSS is easily acquired in receiving sensors, we collect RSS as sensing parameters. The RSS at m_j is expressed as

$$P_r(s_i, m_j) = P_t(s_i) - L(s_i, m_j), \quad (2)$$

where $P_r(s_i)$ means the transmission power of the signal source at s_i. For the whole target area, the RSS from each receiving point is collected into the set $H_{\text{full}} = \{P_r(s_i, m_j)\}$. However, it is impractical to deploy sensors densely. This paper focuses on the localization solutions by learning the propagation features from sparse sensing data. The following is the generation of sparse sensing datasets for deep learning.

2.2. Dataset Preparation. Intelligent cognition of the mapping between sparse sensing data and signal positions is driven by the calculation of electromagnetic data with geographic model. Toward this end, there are three steps to generate the datasets.

Step 1. Calculate the H_{full} of the target area with fixed signal source position s_i, and then extract RSS for sparse sensors from H_{full}. The set of sensor positions is expressed as $V = \{v_1, v_2, \cdots, v_N\}$ with $N \ll Q$ and $V \subseteq M$. Therefore, the RSS of sensors is

$$P_r(V) = P_r(s_i, m_j), m_j \in V. \quad (3)$$

Step 2. Convert RSS data of sensors into an image matrix where the sensing data are combined with sensor positions. Corresponding to the size of target area, the image matrix is established as

$$I(x, y) = \begin{cases} P_r, & (x, y) \in V, \\ 0, & (x, y) \notin V. \end{cases} \quad (4)$$

There are RSS values at the sensor positions, while others are zeros. The DNN is aimed at capturing both geographic and propagation features from the input matrix I.

Step 3. Predict the signal source position in terms of probability distribution. Different from conventional datasets labeled with two-dimensional coordinates, the labels in our dataset are expressed in the image matrix Y with the same size as I. The maximum probability value in Y is 1, located at the true position, and others are zeros. The output $\widehat{Y} = f_{\text{DNN}}(I)$ of the neural network is expected as a probability matrix where the higher the probability is predicated, the closer to the true position.

Then, a fusion dataset is obtained with fixed sensor positions. Its size depends on the number of samples with varied signal positions. In the training dataset, positions of signal sources traverse the target area by gird. In the validation and testing dataset, positions of signal sources are randomly generated in the target area. In addition, different datasets are created by changing the positions and number of sensors. Consequently, there are two advantages of the proposed fusion datasets.

(1) In content, the fusion datasets are collected based on the electromagnetic propagation calculation fused with geographic models, and the datasets are more effective than other simulation datasets based on the empirical or statistical propagation models. Hence, it is more persuasive to verify the validity of proposed localization method by applying the effective datasets

(2) On structure, the fusion datasets are framed in image matrices including position features of sensors and signal sources, highlighting the propagation characteristic from different positions. Besides, compared with the datasets based on the practical measurement, it is more flexible for the proposed datasets to adjust sensor distribution to satisfy the requirements of researches

In conclusion, the proposed method can be used to generate authentic datasets flexibly and efficiently with accessible geographic models.

3. Deep Learning-Based Localization Method

The deep learning-based localization works in two stages, i.e., training stage and localization stage. During the training stage, DNN is driven by fusion datasets to minimize the errors between the predicated matrix $\widehat{Y}$ and the label Y, in order to boost the cognition for the NLOS propagation channels. During the localization stage, the sensing image

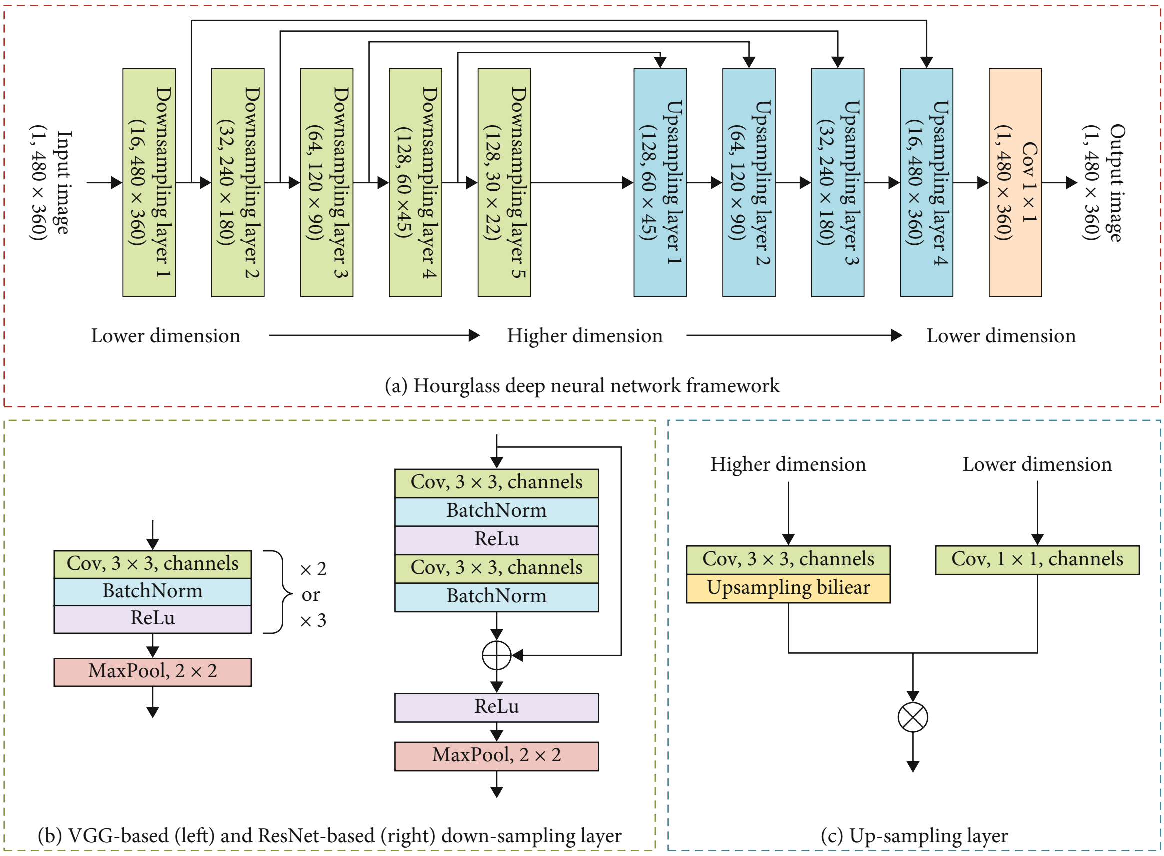


Figure 1: The structure and core modules of hourglass deep neural network.

matrix is sent into the trained network to output the probability matrix. The pixel with the peak value in the output matrix is predicted as the position of single signal source.

In order to improve localization under the sparse layout of sensors, an hourglass architecture is applied to capture multiscale features from the sparse input I. In addition, a weighted loss function is proposed to help the hourglass network pay more attention to the position feature from the sparse label Y.

3.1. The Hourglass Architecture in DNN. Most neural networks extract and learn the features via downsampling, whereas hourglass networks [32] carry out the upsampling after downsampling. In this paper, the upsampling is used to learn the position probability distribution from the multidimensional features extracted by the downsampling. Specifically, the higher-dimensional and lower-dimensional features are combined in the upsampling layers to generate the probability matrix, and the output is restored into the original size of input image gradually. As shown in Figure 1, the hourglass network in this paper contains five downsampling layers and four upsampling layers.

During the downsampling in this paper, the core modules of VGG [33] and ResNet [34] are taken as the feature extractors, respectively. VGG is a convolutional deep neural network proposed in 2014. Each convolutional module consists of a convolutional layer with the kennel of 3×3, a batch normalization layer, and an activation layer with the ReLu function, which leads a deep capture for features. ResNet-based extractor is motivated by the residual module, which has a skip connection between the input and output of a stack of two convolutional layers. And then, the fusion feature is activated by the ReLu function. Moreover, thanks to the max pooling layer after the convolutional module or the residual module, the feature maps become smaller from the lower dimension to the higher dimension.

During the upsampling, the higher-dimensional features and the lower-dimensional features are, respectively, unified in channels by the 3×3 convolutional layer and the 1×1 convolutional layer at first. Next, the bilinear interpolation is adopted to scale up the higher-dimensional feature maps. Then, multiply the higher-dimensional feature maps and the lower-dimensional to realize the fusion of different feature dimensions. Progressively, the size of data turns back to the same as input. Finally, an output matrix $\hat{Y}$ is obtained through a convolutional layer.

3.2. The Proposed Weighted Loss Function. The training stage is intended to minimize the loss between the output $\hat{Y}$ and the label Y. The loss function is expressed in mean square error (MSE) as follows:

$$L_{\mathrm{MSE}} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{Y}_i - Y_i \right)^2. \quad (5)$$

The label Y is extremely sparse, which is intractable for the hourglass network to concentrate on the features around the signal position. Therefore, a weighted loss function is proposed in this section. Without any prior information of illegal signal sources, it is difficult to calculate the weight at each pixel of the matrix precisely. However, we cannot overlook the fact that a higher number of weight should be assigned if it is closer to the true position. Assume that the weight distribution is centered on the label position and decreases around with the two-dimensional normal distribution probability density function. The weight matrix W with the same size as the input and output is expressed as

$$W_i(x, y) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[(x - x_i)^2 + (y - y_i)^2\right]\right\}, \quad (6)$$

where (x_i, y_i) means the coordinate of the illegal signal source s_i, as well as the weighted center. Weighted radius is set as $R = 3\sigma$, which means there is a 99.7 percent chance that the signal source is located in the weighted area. More specifically, the hourglass network expands the attention on the position feature from single pixel to an area of the radius R via the weighted method. The weighted loss function is determined as follows:

$$L_{\mathrm{WMSE}} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{W}_i\left(\widehat{\mathrm{Y}}_i - \mathrm{Y}_i\right)^2. \quad (7)$$

The RMSProp optimizer and backpropagation algorithm are used for training until the loss falls below a threshold or the iteration reaches to a certain number.

In the stage of localization, the RSS data are collected from sensors and converted into the image data. Successively, the trained hourglass network is expected to predict the probability matrix from the sensing image. For the localization task with single signal source, the predicted position is calculated as

$$\widehat{s}_i(\widehat{x}_i, \widehat{y}_i) = \underset{\widehat{x}_i, \widehat{y}_i}{\operatorname{argmax}} \left(\widehat{Y}\right). \quad (8)$$

4. Results and Discussion

4.1. Simulation Setup. As is shown in Figure 2, the geographic model of Tsinghua University is established for the following simulations. A study area with the size of 480 m × 360 m is marked in the red box. Accordingly, the input and output data are 480 ×360-sized image matrices. Based on the geographic model, the RSS set H_{full} is obtained by a software called WinProp [35] which supports the calculation for DPM. The calculation resolution is set to 1 m, representing the interval between receiving points in H_{full} is 1 m. The frequency of signal source is set to 1800 MHz, and the transmission power is 43 dBm, which are unknown for DNN. All simulation parameters are set in detail according to Table 1. Figure 3(a) shows the RSS distribution in the whole study area, which reveals the nonuniform attenuation of radio wave propagation to the surrounding due to the blocking of buildings etc. In Figure 3(b) about the LOS/NLOS distribution, the NLOS propagation paths account for a huge proportion in the urban scenario.

Figure 2: The established three-dimensional geographic model of Tsinghua University (grey, green, and blue objects, respectively, represent buildings, vegetation, and rivers).

Four groups of datasets are set up with varied number of sensors, i.e., 48, 12, 8, and 4. The sensors are uniformly deployed as shown in Figure 4. Compared with most datasets in deep learning-based localization methods [26, 27], the sensor density in our datasets is extremely sparse. For the training datasets, the traversal gird of signal source positions is set into a size of 3 m × 3 m, generating 19200 samples to learn the propagation features of the whole study area. The learning rate is set to 0.001, and the training epoch is 20. To evaluate the performance of the localization in the testing stage, the mean positioning error of the $n_s = 2400$ samples are calculated by

$$e = \frac{1}{n_s}\sum_{i}^{n_s} \sqrt{(x_i - \widehat{x}_i)^2 + (y_i - \widehat{y}_i)^2}. \quad (9)$$

Additionally, in this paper, PyTorch is the basis of the deep learning algorithm framework. All simulations are implemented on an RTX2080ti GPU with 96 GB of RAM.

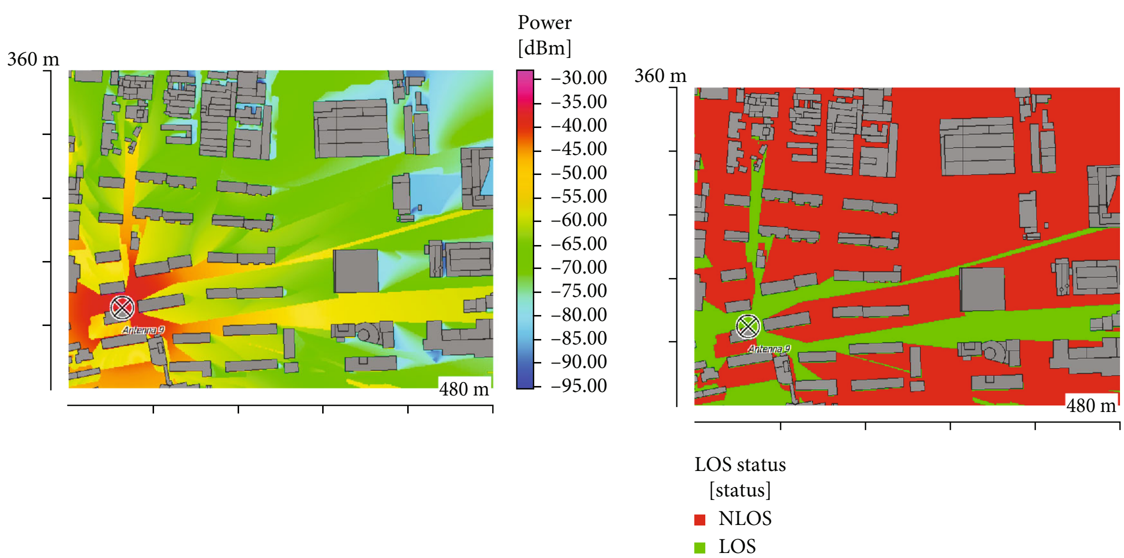
4.2. Results and Evaluation

4.2.1. Comparison of VGG-Based and ResNet-Based Extractor. In this paper, VGG and ResNet are, respectively, applied as feature extractors in downsampling of the hourglass network. From Figure 5, the two extractors achieve satisfactory localization results with the errors of about 2 m when the total number of sensors is 48. With the reduction of sensors, the positioning errors present an upward trend, whereas the ResNet-based extractor suffers less severe degradation than VGG-based extractor. It is indicated that the residual module plays an active part to extract features from sparse input image, which benefits from the fusion of multiscale features though the skip connections. Nonetheless, when there are only 4 sensors, ResNet-based extractor performs worse.

4.2.2. Evaluation of the Weighted Loss Function. To alleviate the hardships of localization from sparse sensing data, the weighted loss function is proposed to enhance the attention on the position features. As the number of sensors decreases,

Table 1: Simulation parameters.

Study area	Size	480 m × 360 m
Buildings	Materials	Concrete
	Thickness of walls	10 cm
	Permittivity	4
	Permeability	1
	Conductivity	0.01 S/m
Water	Permittivity	81
	Permeability	1
	Conductivity	0.5 S/m
Vegetation	Additional loss	10 dB
	Additional attenuation or rays	0.05 dB/m
Signal source	Frequency	1800 MHz
	Power	43 dB
DPM	Resolution	1 m
	Interaction loss	11 dB
	Building penetration loss	20 dB
	LOS path loss exponent	Before/after breakpoint: 2.3/3.3
	NLOS path loss exponent	Before/after breakpoint: 2.5/3.6
	Wave guiding weighting factor	1



(a) (b)

Figure 3: Visual implementation of (a) RSS distribution and (b) LOS/NLOS distribution in the study area of 480 m × 360 m.

the hourglass deep neural network requires wider attention on the position features, leading to a larger weighted radius. However, if the weighted radius is too large, the network has to pay attention to more extra features in the whole weighted area and reduce the attention to the true position. There is a balance between larger weighted area and the better accuracy. Varied from 65 m to 145 m with the interval of 20 m, an appropriate weighted radius is determined via several purposeful setting and numerous tests. Finally, considering the overall positioning performance with different number of sensors, the weighted radius is set to 105 m.

Figure 6 gives the effect of the weighted loss function with a 105 m radius on the localization results, which reveals that the proposed weighted loss method provides a significant enhancement over the original loss function. The localization errors are limited in 10 m with a total of 8 sensors. And the weighted loss function improves the performance of localization more than 50% with a total of 4 sensors.

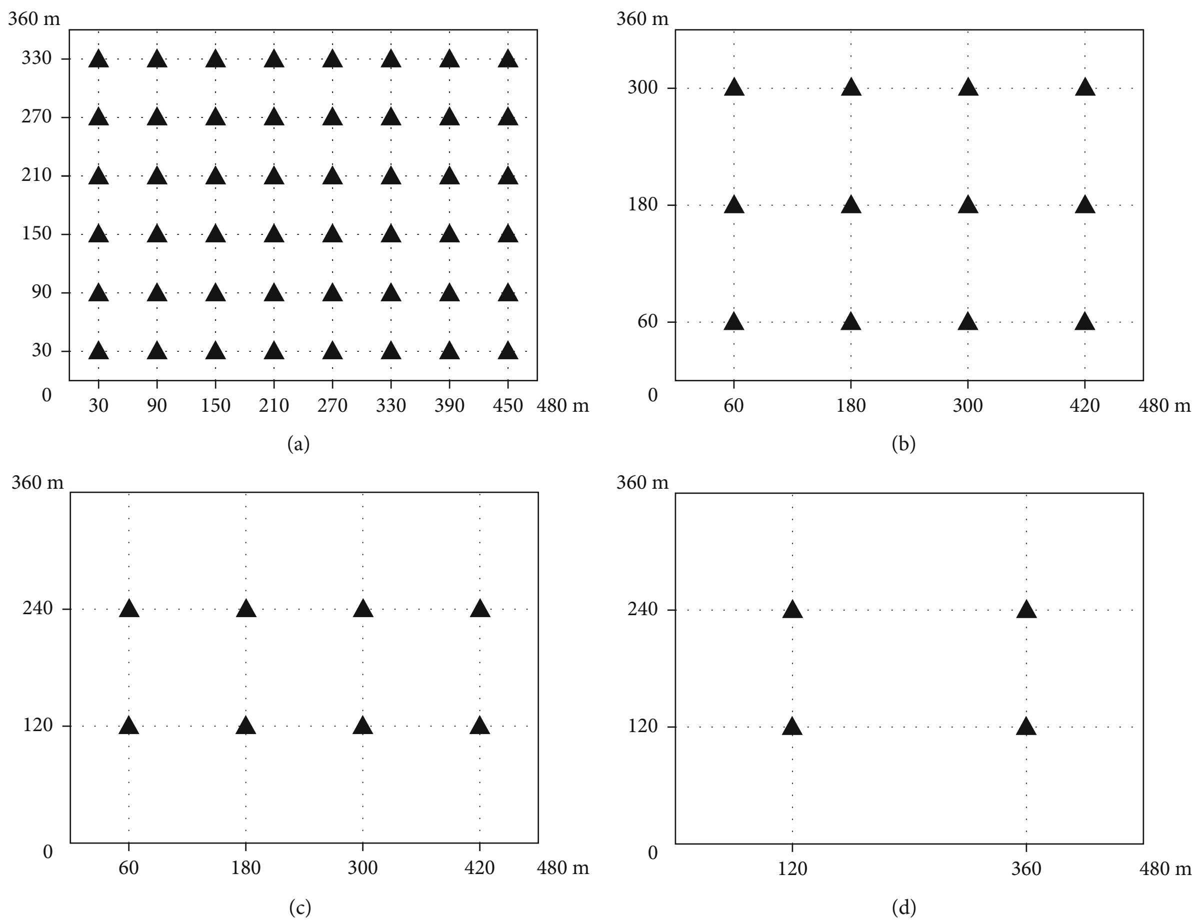


Figure 4: Visual implementation of the distribution of (a) 48 sensors, (b) 12 sensors, (c) 8 sensors, and (d) 4 sensors (which are represented as black triangles).

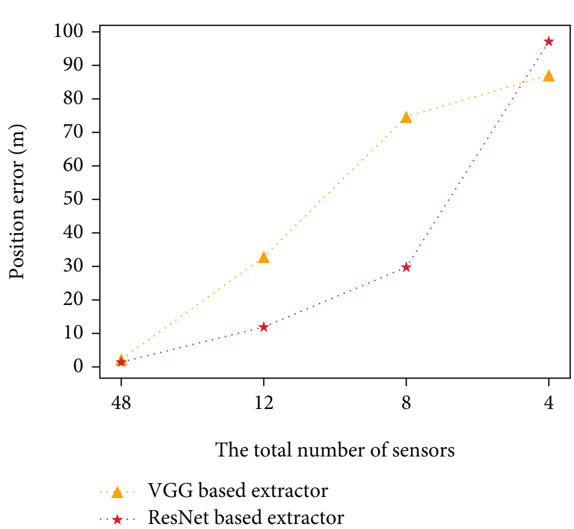


Figure 5: Comparative localization results of VGG-based and ResNet-based extractor with different number of sensors.

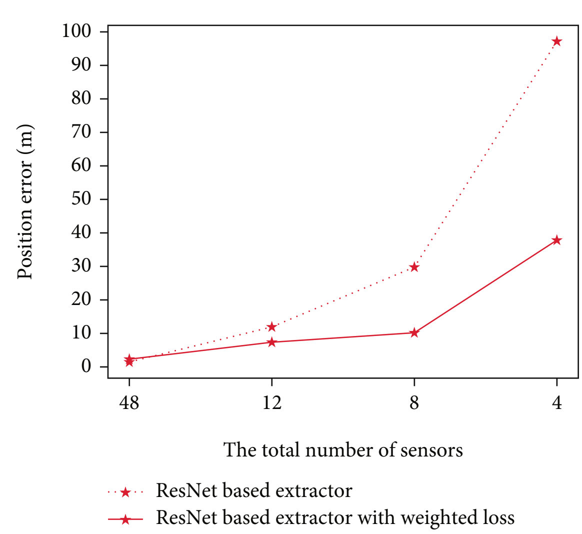


Figure 6: Effect of the weighted loss function on the localization results.

4.2.3. Analysis of LOS/NLOS Propagation. It is rarely possible to count the number of LOS or NLOS sensors without geographic information, so that the sensors with NLOS propagation paths are generally assumed in most simulations of other researches (e.g., [16]). On the other hands, dense layouts of hundreds of sensors in traditional deep learning-

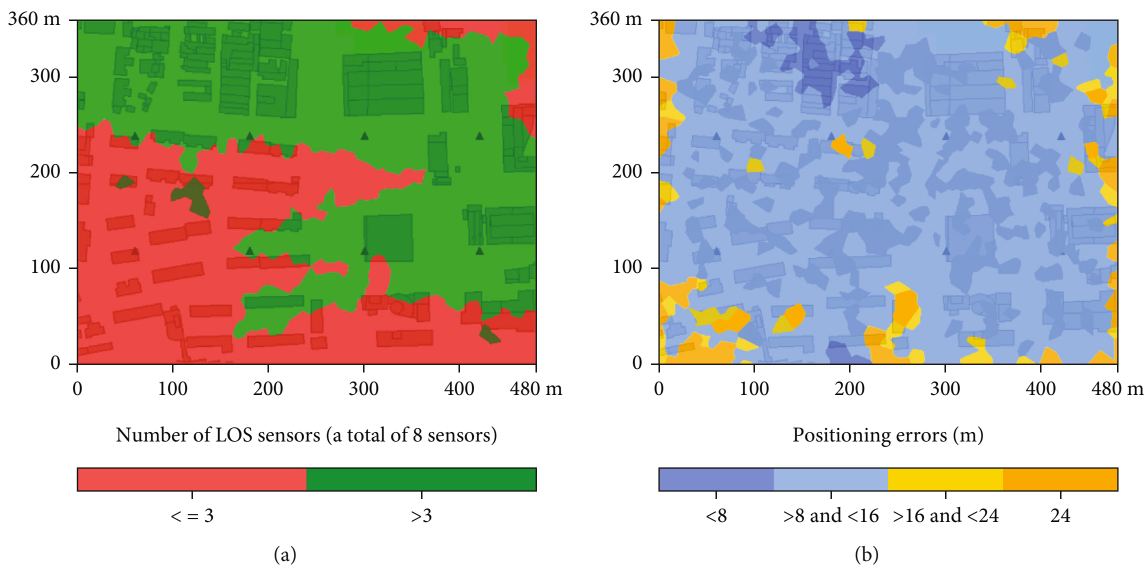


FIGURE 7: Visual implementation of (a) LOS/NLOS distribution and (b) positioning errors with a total of 8 sensors (which are represented as black triangles).

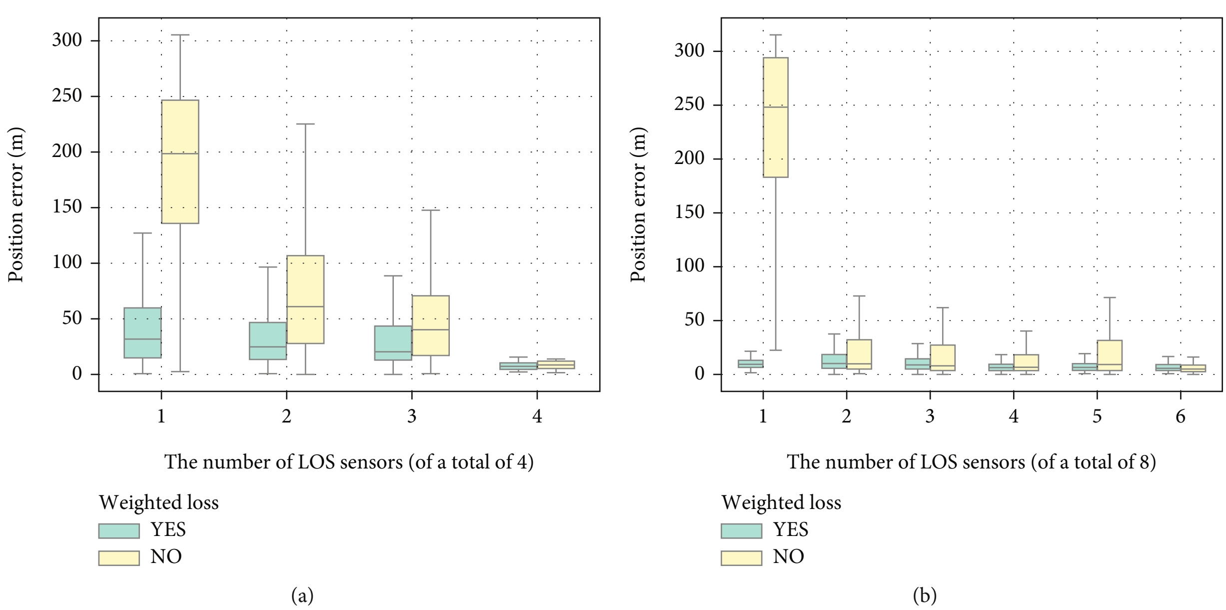


FIGURE 8: The positioning errors with different number of LOS sensors. (a) The total number of sensors is 4. (b) The total number of sensors is 8.

based localization methods inevitably provide sufficient LOS paths for accurate localization. Differently, this paper applies the geographic model to offer LOS and NLOS information and focuses on the localization with sparse distribution of sensors. As a result, the LOS and NLOS sensors are distinguished to evaluate their influences on the localization.

As a qualitative analysis, Figure 7 is the visual implementation of LOS/NLOS distribution and positioning errors against a backdrop of urban buildings. In Figure 7(a), when an illegal signal source is located in the red area, there are less than 4 LOS sensors with a total of 8 sensors, which means it suffers from severe NLOS propagation. In contrast, Figure 7(b) shows that localization results calculated by the proposed method are less affected by NLOS propagation in the corresponding area, indicating that the NLOS propagation features are well captured for robust localization. And the worse performance at the edge of the study area can be interpreted by the poor sensing.

As a quantitative analysis, Figure 8 gives the positioning errors with different number of LOS sensors, representing that the more LOS sensors there are, the better performance the proposed method offers. With a total of 4 sensors, Figure 8(a) shows that the error fluctuations of the localization method with weighted loss function decrease 85% as the number of LOS sensors increases from 1 to 4. In addition, with a total of 8 sensors in Figure 8(b), the weighted method

Table 2: Comparative results of different algorithms.

Algorithms	Positioning errors (m) with varied number of sensors 48	12	8	4
DeepMTL	19.00	203.08	223.53	222.59
U-net	5.48	35.95	100.69	140.13
VGG-based proposed method	2.36	9.92	16.40	50.87
ResNet-based proposed method	2.28	7.34	10.16	37.81

achieves the robust localization performance of median errors below 10 m, especially as the original method fails to localize with single LOS sensor.

4.2.4. Comparison with Other Algorithms. The proposed method is compared with DeepMTL localization [27] and another network structure, i.e., U-net [36]. DeepMTL primarily converts the RSS value into image data and predicts the positions of multiple transmitters by CNN. In this paper, apply DeepMTL for localization of single signal source to evaluate its performance under the condition of NLOS propagation and sparse sensor distribution. Besides, a typical U-net that consists of four downsampling and four upsampling layers is considered for comparisons.

Table 2 shows the comparative results. DeepMTL performs worst due to the sparse layouts of sensors. Its original performance is based on the dense sensors and the ideal or statistical propagation data which are easy to learn the localization features. Driven by our datasets that is closer to the practical electromagnetic environment, its localization errors of single signal source are up to 200 meters, illustrating that DeepMTL is not suitable for the severe NLOS propagation conditions. On the other hand, the VGG-based and ResNet-based hourglass networks with weighted loss function have a 63.7% and 73.0% improvement over U-net with a total of 4 sensors, respectively. The comparisons express the fusion of multiscale features in hourglass network which is more efficient to predict probability distributions of positions than U-net, and the weighted loss function improves the attention on true positions. Therefore, it is concluded that the proposed method is more qualified for localization in large-scale urban scenarios.

5. Conclusions

For the challenges of severe NLOS propagation and sparse layouts of sensors in urban scenarios, this paper proposes a deep learning-based localization method which learns the NLOS propagation features with the aid of urban geographic model. The proposed method fully considers the fusion of geographic information and electromagnetic wave propagation. The reliable urban propagation model prepares the datasets close to practical electromagnetic environment and motivates neural networks. Moreover, the improved hourglass networks with two typical feature extractors, i.e., VGG and ResNet, are trained for localization. In comparative simulations, the ResNet-based hourglass network outperforms the VGG-based, but huge positioning errors still exist when there are only 4 sensors. With the assistance of the proposed weighted loss function, the localization performance is enhanced by over 50%. Furthermore, geographic information supports for the distinction between LOS and NLOS paths in order to evaluate the robustness of the method with different number of LOS sensors. Compared with other deep learning-based algorithms, the proposed method keeps obviously reliable and superior performance as the total number of sensors decreases. The results present the validity and robustness of the proposed deep learning-based localization method under the severe NLOS propagation and sparse sensing conditions in urban scenarios.

Data Availability

The original data used in this work is generated by simulation via WinProp. The method of dataset generation is included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (grant numbers 61971439 and 61702543).

References

[1] A. Haniz, G. K. Tran, K. Saito et al., “A novel phase-difference fingerprinting technique for localization of unknown emitters,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8445–8457, 2017.

[2] B. Kaur and D. Prashar, *Analysis and Comparison of Localization Approaches in WSN: A Review*, Springer, Singapore, 2019.

[3] I. Javed, X. Tang, M. A. Saleem, M. U. Sarwar, M. Tariq, and C. S. Shivachi, “3D localization for mobile node in wireless sensor network,” *Wireless Communications and Mobile Computing*, vol. 2022, 12 pages, 2022.

[4] J. Pospisil, R. Fujdiak, and K. Mikhaylov, “Investigation of the performance of TDoA-based localization over LoRaWAN in theory and practice,” *Sensors*, vol. 20, no. 19, article 5464, 2020.

[5] Y. Chen, Z. Yao, and P. Zheng, “A novel method for asynchronous time-of-arrival-based source localization: algorithms, performance and complexity,” *Sensors*, vol. 20, no. 12, p. 3466, 2020.

[6] C. Xu, Z. Wang, Y. Wang, Z. Wang, and L. Yu, “Three passive TDOA-AOA receivers-based flying-UAV positioning in

extreme environments," *IEEE Sensors Journal*, vol. 20, no. 16, pp. 9589–9595, 2020.

[7] Y. Li, F. Shu, B. Shi, X. Cheng, Y. Song, and J. Wang, "Enhanced RSS-based UAV localization via trajectory and multi-base stations," *IEEE Communications Letters*, vol. 25, no. 6, pp. 1881–1885, 2021.

[8] S. Tomic, M. Beko, and R. Dinis, "3-D target localization in wireless sensor networks using RSS and AoA measurements," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3197–3210, 2017.

[9] L. Cheng, M. Xue, Y. Wang, Y. Wang, and Y. Bi, "A robust tracking algorithm based on modified generalized probability data association for wireless sensor network," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 2, pp. 2136–2146, 2022.

[10] J. Hua, Y. Yin, W. Lu, Z. Yu, and F. Li, "NLOS identification and positioning algorithm based on localization residual in wireless sensor networks," *Sensors*, vol. 18, no. 9, p. 2991, 2018.

[11] C. L. Sang, B. Steinhagen, J. D. Homburg, M. Adams, M. Hesse, and U. Rückert, "Identification of NLOS and multi-path conditions in UWB localization using machine learning methods," *Applied Sciences*, vol. 10, no. 11, p. 3980, 2020.

[12] Y. Yuan, F. Shen, and X. Li, "GPS multipath and NLOS mitigation for relative positioning in urban environments," *Aerospace Science and Technology*, vol. 107, article 106315, 2020.

[13] J. Bordoy, C. Schindelhauer, R. Zhang, F. Hoflinger, and L. M. Reindl, "Robust extended Kalman filter for NLOS mitigation and sensor data fusion," in *2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pp. 117–120, Kauai, HI, USA, March 2017.

[14] S. Tomic and M. Beko, "A robust NLOS bias mitigation technique for RSS-TOA-based target localization," *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 64–68, 2019.

[15] C. Geng, X. Yuan, and H. Huang, "Exploiting channel correlations for NLOS ToA localization with multivariate Gaussian mixture models," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 70–73, 2020.

[16] Y. Sun, S. Yang, G. Wang, and H. Chen, "Robust RSS-based source localization with unknown model parameters in mixed LOS/NLOS environments," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3926–3931, 2021.

[17] S. Miura, L.-T. Hsu, F. Chen, and S. Kamijo, "GPS error correction with pseudorange evaluation using three-dimensional maps," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3104–3115, 2015.

[18] D. Acharya, M. Ramezani, K. Khoshelham, and S. Winter, "BIM-Tracker: a model-based visual tracking approach for indoor localisation using a 3D building model," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 157–171, 2019.

[19] G. Fuqiang, S. Valaee, K. Khoshelham, J. Shang, and R. Zhang, "Landmark graph-based indoor localization," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8343–8355, 2020.

[20] M. Zhou, Y. Lin, N. Zhao, Q. Jiang, X. Yang, and Z. Tian, "Indoor WLAN intelligent target intrusion sensing using ray-aided generative adversarial network," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 1, pp. 61–73, 2020.

[21] G. Zhang, H.-F. Ng, W. Wen, and L.-T. Hsu, "3D mapping database aided GNSS based collaborative positioning using factor graph optimization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6175–6187, 2021.

[22] F. Perez-Cruz, P. M. Olmos, M. M. Zhang, and H. Huang, "Probabilistic time of arrival localization," *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1683–1687, 2019.

[23] I. Nirmal, A. Khamis, M. Hassan, H. Wen, and X. Zhu, "Deep learning for radio-based human sensing: recent advances and future directions," *IEEE Communications Surveys Tutorials*, vol. 23, no. 2, pp. 995–1019, 2021.

[24] Y. Liu, C. Lyu, X. Liu, and Z. Liu, "Automatic feature engineering for bus passenger flow prediction based on modular convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2349–2358, 2021.

[25] A. Poulose and D. S. Han, "UWB indoor localization using deep learning LSTM networks," *Applied Sciences*, vol. 10, no. 18, p. 6290, 2020.

[26] M. Lin, Y. Huang, B. Li, and Z. Huang, "Heatmap-basedmultiple co-channel transmitter localization with fully convolutional network," in *2021 International Applied Computational Electromagnetics Society (ACES-China) Symposium*, pp. 1-2, Chengdu, China, 2021.

[27] C. Zhan, M. Ghaderibaneh, P. Sahu, and H. Gupta, "DeepMTL: deep learning based multiple transmitter localization," in *2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 41–50, Pisa, Italy, June 2021.

[28] L. M. Correia, "A view of the COST 231-Bertoni-Ikegami model," in *2009 3rd European Conference on Antennas and Propagation*, pp. 1681–1685, Berlin, Germany, 2009.

[29] B. Friebel, M. Schweins, N. Dreyer, and T. Kurner, "Simulation of GPS localisation based on ray tracing," *Advances in Radio Science*, vol. 19, pp. 85–92, 2021.

[30] R. Wahl, G. Wolfle, P. Wertz, P. Wildbolz, and F. Landstorfer, "Dominant path prediction model for urban scenarios," in *14th IST Mobile and Wireless Communications Summit*, pp. 1–5, Dresden, Germany, 2005.

[31] A. Aschrafi, P. Wertz, M. Layh, F. M. Landstorfer, G. Wolfle, and R. Wahl, "Impact of building database accuracy on predictions with wave propagation models in urban scenarios," in *2006 IEEE 63rd Vehicular Technology Conference*, vol. 6, pp. 2681–2685, Melbourne, VIC, Australia, 2006.

[32] H. Law and J. Deng, "CornerNet: detecting objects as paired keypoints," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 734–750, Munich, Germany, September 2018.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[35] R. Hoppe, G. Wolfle, and U. Jakobus, "Wave propagation and radio network planning software. WinProp added to the electromagnetic solver package FEKO," in *2017 International Applied Computational. Electromagnetics Society Symposium - Italy (ACES)*, pp. 1-2, Firenze, Italy, March 2017.

[36] Z. Wang, Y. Zhou, F. Wang, S. Wang, and X. Zhiyu, "SDGH-Net: ship detection in optical remote sensing images based on Gaussian heatmap regression," *Remote Sensing*, vol. 13, no. 3, p. 499, 2021.