A lightweight CIR-based CNN with MLP for NLOS/LOS identification in a UWB positioning system

Minghao Si, Yunjia Wang, Harun Siljak, CheeKiat Seow, Hongchao Yang

***Abstract*—Implementing line-of-sight (LOS) and none-line-of-sight (NLOS) identification in ultra-wideband (UWB) systems is crucial. Convolutional neural network (CNN) based identification methods can extract higher-level features automatically, but they are based on channel impulse response (CIR)-turned image ingested features that impose calculation complexity and do not make use of manual features due to the data inundation risk. In this letter, we propose a novel multilayer perceptron (MLP)-based LOS/NLOS identification algorithm that can utilize both manually extracted features and feature from CNN based on raw CIR inputs with only 7.39% calculation complexity as compared to the traditional image-based CNN. Three experiments, at a teaching building, an office, and an underground mine, were conducted to verify the proposed method's performance. Our proposed features are conducive to LOS/NLOS identification, especially the proposed raw CIR-based feature from the CNN, achieving 26.9% improvement over existing manual features. Furthermore, the proposed method outperformed the traditional image-based CNN with an improvement of 44.16%.**

***Index Terms*—UWB, LOS/NLOS identification, CNN, MLP.**

I. Introduction

LINE-of-sight (LOS)/none-line-of-sight (NLOS) identification is crucial for the ultra-wideband (UWB) positioning system [1]. Many LOS/NLOS identification methods have been proposed to improve accuracy. Most of them are based on Gaussian, support vector machines (SVMs) and K-Nearest Neighbours (KNNs) [2]. However, these methods require manual feature extraction and considerable manpower to adjust suitable feature parameters.

Compared with the aforementioned methodologies, those based on convolutional neural networks (CNNs) can achieve high accuracy and do not require manual features extraction or parameters adjustment [3], [4]; hence, they are time and labour efficient. In a previous study [5], a CNN was trained to analyse an image-based channel impulse response (CIR) matrix to identify the LOS/NLOS conditions. In another study [6], the author proposed a LOS/NLOS identification method based on image-based CIR using Morlet wavelet transform and convolutional neural networks (MWT-CNN). However, the inputs of these methods are images, whose sizes directly determine the accuracy of these methods. Thus, a high-accuracy image-based CNN will result in major computational complexity. Furthermore, in other studies [7], [8], we can infer that manually extracted features are beneficial in the accuracy improvement of LOS/NLOS identification, but CNN-based methods cannot directly use the raw CIR and manually extracted features simultaneously because the lengths of the raw CIR (approximately 1,000) and the manually extracted features are different. Direct combination will result in a small number of useful manual features being obfuscated under the large CIR dataset. In conclusion, there are challenges in reducing the computational complexity for CNN and in combining CNN with manual features to achieve high accuracy without data inundation risk. In this letter, we propose a LOS/NLOS identification algorithm for UWB positioning systems, considering the above challenges. The main contributions of this letter include the following:

- A raw CIR-based CNN is proposed instead of a traditional image-based CNN to extract feature. Through comparison, the calculation complexity of the CIR-based CNN is only 7.39% of that of the 64×64 image-based CNN.
- Through the analysis of the noise threshold multiplier (NTM), first path, ranging results and CIR waveform signal distribution of UWB, three additional manual features of the CIR for LOS/NLOS identification have been proposed that can improve the identification accuracy.
- A multilayer perceptron (MLP)-based LOS/NLOS identification algorithm is proposed to fuse the feature from the CNN and the six manual features, including three proposed features, to improve the accuracy. This assimilation provides effective utilization of both manual features and the CIR without data inundation risk.

To the best of our knowledge, there is no existing literature that has developed raw CIR-based CNNs and the fusion of CNN features and manual features.

II. Theoretical Analysis and Proposed Method

A. Analysis of calculational complexity and data inundation risk

Generally, there are three different layers in a CNN, namely convolution layers (CLs), pooling layers (PLs) and fully connected layers (FCLs). Traditional CNN-based identification methods convert CIR into images followed by conversion into matrix data as training input. However, given the length of the raw CIR, image-based CNN is not suitable for use

Minghao Si, Yunjia Wang, Hongchao Yang are with the Key Laboratory for Land Environment and Disaster Monitoring, China University of Mining and Technology, Xuzhou, 221116, China (e-mail:hmsi@cumt.edu.cn;wyjc411@.163.com;hongchao_yang@cumt.edu.cn).

Chee Kiat Seow is with the School of Computing Science, University of Glasgow, Glasgow G12 8RZU.K (e-mail:Cheekiat.Seow@glasgow.ac.uk).

Harun Siljak is with the School of Engineering, Trinity College Dublin, the University of Dublin, Dublin 2, Ireland (e-mail:harun.siljak@tcd.ie).

because this would entail much computing inefficiency. As shown in Fig. 1, a CIR 64×64 image and a raw CIR data are selected as the input of the CNN. The number of layers is the same, while the kernel and pooling dimensions of the image-based CNN expand from one dimension to two dimensions. To estimate the computational complexity, we count the number of operations in the convolution and pooling layers, and then sum them up. For the convolution layer, we multiply the number of convolution operations by the total number of additions and multiplications involved in each convolution. As for the pooling layer, we calculate the total number of pooling operations and then multiply it by the total number of additions and divisions involved in each pooling operation. The image-based CNN executed a total of 98,516 operations, while the raw CIR executed only 7,289 operations. Therefore, the computational complexity of the CIR-based CNN is only 7.39% of that of the image-based CNN. To reduce computational complexity, we propose using raw CIR to train the CNN.

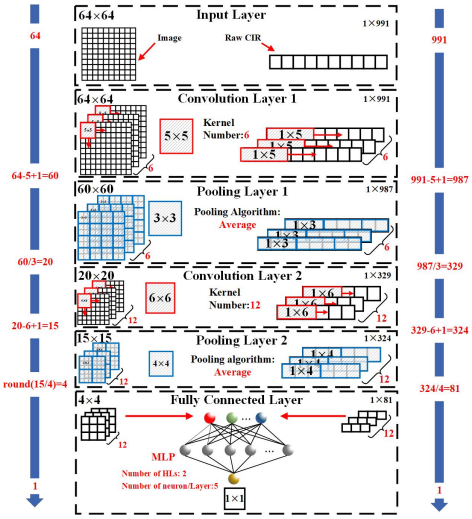

Fig. 1. The comparison of image-based CNN and raw CIR-based CNN

However, the length of a normal UWB CIR is approximately 1,000, while the length of manual features is less than 10; thus, a simple combination of raw CIR and manual features at the input layer will result in serious data inundation. As shown in Fig. 1, after two convolutions and pooling, the length of the raw CIR data is compressed to 81. As seen, if the proposed and existing manual features (six in this case) are simply combined with the raw CIR (1×991) at the input layer, these manual features will be obfuscated. This means that the obfuscated manual features make no contribution in the accuracy improvement.

B. Proposed Method

The LOS/NLOS identification processes are divided into offline training and online deployment components, as shown in Fig. 2. In offline training, the input features arise from CNN extraction and manual features to train the MLP identification model. During the online phase, the MLP can identify LOS/NLOS signals in real time. The features extraction and combination details are introduced below.

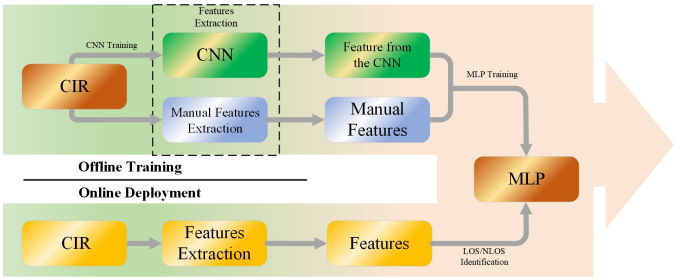

Fig. 2. Methodology framework

1) Features Extraction: In total, seven features were extracted for MLP training, including the new CIR-based feature from the CNN, three proposed manual features and three existing valid features.

a) Feature from the CNN: Figure 3 shows the raw CIR and the resulting 1x81 and 1x7 feature maps after convolution and pooling. The convolution increases the difference between LOS and NLOS, indicating that data compression is effective. Therefore, after inputting them to the FCL, it will output a LOS probability that can well distinguish between LOS and NLOS, which is considered as a feature from the CNN (the 1x1 MLP output in Fig.1) and will be subsequently combined with the manual features.

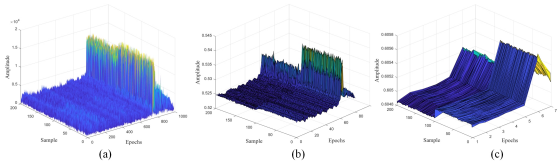

Fig. 3. Comparison between pre-convolution and post-convolution. (a) CIR of LOS/NLOS before the CNN, (b) CIR after Convolution and Pooling to 1x81, (c) CIR after Convolution and Pooling to 1x7

b) Proposed additional features: In addition to the feature from the CNN, we proposed three additional features that would perform differently in the LOS and NLOS environments.

• First Path Strength Difference ($FPSD$). In NLOS scenarios, obstacles may weaken the strength of the true first path such that it does not exceed the first path identification threshold L_1. Consequently, a stronger multipath signal may be incorrectly identified as the first path. The difference between the signal strength of the reported first path (RFP) and the potential true first path differs in the LOS and NLOS environments and can be used for identification. Therefore, we recalculated the identification threshold and then calculated the difference as $FPSD = RFP - \underset{r(t_i)}{argmin}(r(t_i) \geq S \cdot NTM)$, where $r(t_i)$ is the signal strength at time t_i within the CIR, S is the standard deviation of the CIR noise level, NTM is the noise threshold multiplier.

• First Path Distance Difference ($FPDD$). The distance d_1 between the tag and the anchor can be estimated with $d_1 = \frac{P_T + G_R + 20\log(c) - FPPL - 20\lg(4\pi f_R)}{10}$, where P_T is the

transmitted power; G_R is the total antenna gains of the anchor and tag antennas; $FPPL$ is the first path power level; and f_R is the centre frequency [9]. Also a ranging result d_2 can be achieved from the tag's report. $FPDD$ can be obtained by calculating the difference between d_1 and d_2. The performance of FPDD is different in the NLOS and LOS environments, as shown in Fig. 4, and this performance difference can be used as a valid feature.

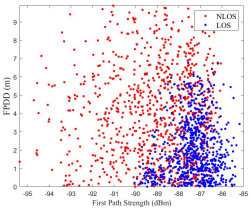

Fig. 4. The distance difference between LOS and NLOS

• Number of Pseudo peaks (NPP). In the NLOS environment, the true first path is often classified as environmental noise, which results in the NPP between L_1 ($L_1 = S \cdot NTM$)and L_2 ($L_2 = 0.6 \cdot L_1$) in the NLOS environment being greater than that in the LOS environment [10]. To verify this, we calculated the average NPP of epochs and their percentage in the total when the NPP between L_1 and L_2 is not equal to 0 in the NLOS and LOS environments, as shown in Table I. The differences in the average and percentage values in the LOS and NLOS environments are obvious, and NPP can, therefore, be a valid distinct feature for LOS/NLOS identification.

TABLE I
NPP OF LOS AND NLOS

	Average	**Percentage** %
NPP of LOS	2.924	3.58
NPP of NLOS	7.287	14.08

c) Other existing features: In addition to the above proposed features, three existing manual features are utilized, namely kurtosis ($k = \frac{1}{N\sigma_r^4}\sum_{i=1}^{N}(|r(t_i)-\mu_r|)^4$), mean excess delay ($\tau_{med} = \frac{1}{\varepsilon}\sum_{i=1}^{N}(t_i|r(t_i)|^2)$) and root-mean-square delay spread ($\tau_{rms} = \frac{1}{\varepsilon}\left[(t_i-\tau_{med})^2|r(t_i)|^2\right]$), where μ_r is the mean of amplitude, σ_r^2 is the variance of CIR, ε is the total energy of $r(t_i)$ and N is the number of discrete samples of the signal waveform [11].

After feature extraction, to determine the correlation between features and avoid overfitting or model performance degradation, it is necessary to investigate the independence between manual features and feature from CNN. We randomly selected 1000 LOS and 1000 NLOS data samples, discretized the 7 features equally, and calculated their marginal probability $p(x_i)$ and joint probability distributions $p(x_i, x_j)$ to obtain their mutual information according to $I(x_i, x_j) = \sum_{x_i \in X_i}\sum_{x_j \in X_j} p(x_i, x_j)\log\frac{p(x_i,x_j)}{p(x_i)p(x_j)}$, i, j is the index of features from 1 to 7 and X_i, X_j is the sets of values of x_i, x_j. Our experimental results showed that the mutual information between τ_{med} and τ_{rms} was the highest, reaching 0.034, while the independence between the other features was below 0.026. Based on these findings and the fact that our manual features capture distinct signal characteristics, we conclude that they are independent. Therefore, the interactions between the features in MLP can be described as simple linear additions, meaning that each feature independently contributes to the final output.

2) Features Combination Based on MLP: The above 7 extracted features are then used as input of a MLP structure, including the feature from the CNN, 3 proposed features and 3 existing features. In the input layer, the length of the feature from the CNN is one, while the length of the manual features is six. Since the length difference is small and MLP can adaptively adjust the weight of features, the combination based on MLP can avoid the data inundation risk.

III. EXPERIMENTS AND ANALYSIS

A. Experimental Setup and Data Collection

Three experimental campaigns, namely, at a teaching building, an office and an underground mine, were conducted to verify the performance of the proposed method. The office environment is relatively simple, with LOS anchor and tag placed in an open space while NLOS anchors were blocked by an iron door and two glass doors, as shown in Fig. 5(b). The teaching building is a classic indoor environment, where the NLOS anchor was deployed in a room while the LOS anchor and tag were placed in the corridor, as shown in Fig. 5(a). In the LOS environment (corridor), there were reflected NLOS signals from the concrete wall during transmission, resulting in the multipath effect. In addition, in this NLOS environment, there were obstacles such as wooden doors, desks, pedestrians, etc., that contributed to the existence of NLOS paths. The underground mine environment is a challenging environment, as it was closed and narrow, with two large metal pipes and many metal supports on both sides of the wall, which easily caused a serious multipath effect. Moreover, the proximity of LOS and NLOS anchors also increased the difficulty in distinguishing, as shown in Fig. 5(c).

The proposed method was evaluated through data collection in three scenarios where the anchors were fixed and the tag was placed at varying distances from the LOS and NLOS anchors. The sampling distance ranged from 0 to 11 m, with a sampling point interval of 1 m. A total of 18,400, 20,000, and 16,000 samples were collected in the teaching building, office, and underground mine, respectively, with a fair equal distribution in LOS and NLOS data. The samples were then split into 70:30 ratios for training and testing data. To ensure consistency and fair performance evaluation, the parameters and configuration of the CNN model used in the three experimental campaigns were kept the same, with two convolution and two pooling layers. The kernel sizes of the convolution layers were 1x5 and 1x6, with 6 and 12 kernel numbers, respectively. The sampling down-scales of the two pooling layers were 3 and 4, respectively. The number of MLP's hidden layers was 2, with 5 neurons in each layer. The MLP utilizes the cross-entropy loss function during training, with a batch size of 20. And the

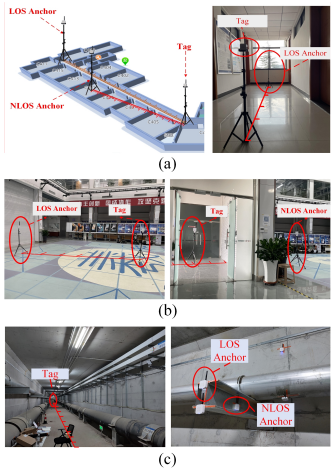

Fig. 5. Deployment scenarios in the underground mine. (a) teaching building (b) office (c) underground mine

training stops when the gradient variation drops below 10^{-4} or the iterative number is over 200.

B. The analysis of proposed features

To verify the effectiveness of the proposed features, we calculated their identification accuracy individually using MLP in three scenarios. The accuracy of a single feature was higher than 60%, as shown in Table II. Notably, the new feature from CNN alone achieved high identification accuracies of 86.73%, 91.56%, and 85.75% in the teaching building, office, and mine, respectively, indicating the effectiveness of feature from CNN. In addition, compared with other features, the ac-curacy of feature from CNN has been significantly improved, especially by 26.9% higher than that of τ_{rms}, demonstrating the novelty and superior performance of our raw CIR feature extraction.

TABLE II
ACCURACY OF SINGLE FEATURE

	Teaching Building %	**Office** %	**Mine** %
Feature from the CNN	86.73	91.56	85.75
FPSD	67.79	72.16	62.55
FPDD	76.31	81.48	66.58
NPP	65.94	71.65	63.89
k	64.30	77.16	60.88
τ_{med}	64.51	68.95	61.65
τ_{rms}	63.54	64.66	60.88

C. The function of feature from the CNN

Using the proposed architecture in Fig. 2, an ablation study was conducted to truly explore the function of feature from CNN, where the proposed method (feature from CNN and 6 manual features) was compared with that without feature from CNN, as shown in Table III. Overall, our proposed architecture achieved more than 90% accuracy in all scenarios. The accuracy of the proposed method in the three scenarios achieve stability with deviation less than 2.39%, as compared to 4.75% that uses only manual features. In addition, using feature from the CNN enhances robustness and environmental resiliency. The improvements in the three scenarios were 4.94%, 1.07%, and 6.43%, respectively, of which the improvement in office was the least, while that in mine was the largest. This is mainly because the feature from the CNN can differentiate LOS and NLOS multipaths more effectively since CNN is supervised learning, and during the training process, it can account for the influence of NLOS multipaths. Compared with office environment, the teaching building and mine environments are more complex with more obstacles. Especially in mine environment, the narrow pathway and metal obstacles exacerbated serious NLOS multipath effects.

TABLE III
COMPARISON OF THE PROPOSED METHOD WITH AND WITHOUT (W/O) FEATURE FROM THE CNN BASED ON PROPOSED ARCHITECTURE IN FIG. 2

Methods	**Teaching Building** %	**Office** %	**Mine** %
Feature from the CNN and manual features	98.85	99.73	97.34
(w/o) Feature from the CNN	93.91	98.66	90.91

D. The robustness and accuracy of the proposed method

Firstly, we conduct a comparison among the proposed method(feature from CNN + manual features), proposed CIR-based CNN (no manual features) and the simple combination based on CNN (raw CIR + manual features), in three scenarios to verify the ability to solve data inundation risk. The CNN parameters were the same. As shown in Table IV, compared with a CIR-based CNN, a simple combination can only improve the accuracy by less than 1%. On the other hand, the proposed method outperforms the simple combination by 11.93%, 8.12% and 11.46% in teaching building, office and mine respectively without data inundation risk.

TABLE IV
COMPARISON OF THE PROPOSED METHOD, CIR-BASED CNN AND SIMPLE COMBINATION

Method	**Teaching Building** %	**Office** %	**Mine** %
Proposed method	98.85	99.73	97.34
Proposed CIR-based CNN	86.73	91.56	85.75
Simple combination based on CNN	86.92	92.11	85.88

We compared the accuracy and time consumption of CNNs based on CIR of different lengths and images of different sizes, as shown in Table V. Notably, when the accuracy of the image-based CNN exceeds 90% for a 256×256 image size, the total inference time is 129 times longer than that of the proposed raw CIR-based CNN. Conversely, when the time consumption is less than 0.1 s, the proposed CIR-based CNN has an accuracy 27.43% higher than the image-based CNN. Considering the requirements of UWB-based applications, the

proposed CIR-based CNNs are more advantageous and better suited to meet real-time requirements. Additionally, compared to the 991×1 CIR, the accuracy differences of the 1000×1 and 1010×1 CIRs are less than 0.22%, and time consumption is almost the same, demonstrating the scalability of the proposed method.

TABLE V
COMPARISON OF IMAGE-BASED AND PROPOSED CIR-BASED CNN

Input	Identification accuracy %	Time consuming s
64x64 image	64.13	0.08
128x128 image	72.33	0.34
256x256 image	97.59	1.29
1x991 raw CIR	91.56	0.01
1×1000 raw CIR	91.73	0.01
1×1010 raw CIR	91.42	0.01

To verify the performance of the proposed method, we proposed two new identification methods based on SVM and KNN (replacing MLP in Fig. 2 with SVM and KNN) with the same inputs as the proposed method. The K value of KNN is 5. Considering the need for real-time identification with a time consumption of less than 0.15s, the 64*64 image-based CNN was chosen as the benchmark. As shown in Fig. 6, although the accuracies of all methods were high in the office, the accuracies of SVM, KNN and image-based CNN dropped to 93.72%, 92.53%, 57.22% in the teaching building and 90.29%, 89.13%, 53.18% in the underground mine. In contrary, our proposed method maintains a much higher accuracy, especially in underground mine with 98.85% in teaching building, 99.73% in office and 97.34% in underground mine, demonstrating the resilience and stability of the proposed method in complex environments. In addition, the time consumption of SVM and the proposed method are approximately 0.13 s and 0.03 s, respectively. In summary, our proposed method outperformed SVM in terms of time consumption and accuracy. Compared with traditional image-based CNN, the accuracy of our method improved by 44.16%.

Finally, we have analysed the performance of the proposed method for different distances between the anchor and the tag in three scenarios, as shown in Fig. 7. The lowest accuracies of the proposed method for distance range of 1 m to 11 m are 97.32%, 99.34% and 96.01% in teaching building, office and mine environment respectively.

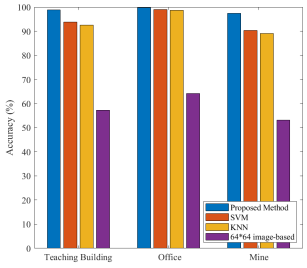

Fig. 6. Comparison of the proposed method and other methods

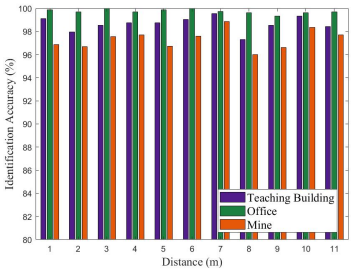

Fig. 7. The performance of the proposed method at different distances

IV. CONCLUSION

This letter proposes a NLOS/LOS identification method for UWB positioning systems. A new feature from CNN and three new manual features are proposed, which are combined with three existing features based on MLP, without the data inundation risk. The calculation complexity of the feature from the CNN is only 7.39% of that of traditional image-based CNN. According to experiments ,our proposed features are conducive to LOS/NLOS identification, especially the proposed feature from the CNN, achieving 26.9% improvement over existing features. Furthermore, the proposed method outperformed the image-based CNN with an improvement of 44.16%.

Our future work will explore the performance of feature from CNN in different environments, and consider setting parallel CNNs to improve accuracy

REFERENCES

[1] K. Paszek, D. Grzechca, and A. Becker, "Design of the uwb positioning system simulator for los/nlos environments," *Sensors*, vol. 21, no. 14, p. 4757, 2021.
[2] M. Si, Y. Wang, C. K. Seow, H. Cao, H. Liu, and L. Huang, "An adaptive weighted wi-fi ftm-based positioning method in an nlos environment," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 472–480, 2021.
[3] H. Zhang, S. Y. Tan, and C. K. Seow, "Toa-based indoor localization and tracking with inaccurate floor plan map via mrmsc-phd filter," *IEEE Sensors Journal*, vol. 19, no. 21, pp. 9869–9882, 2019.
[4] D. T. A. Nguyen, H.-G. Lee, E.-R. Jeong, H. L. Lee, and J. Joung, "Deep learning-based localization for uwb systems," *Electronics*, vol. 9, no. 10, pp. 2196–2200, 2020.
[5] T. Zeng, Y. Chang, Q. Zhang, M. Hu, and J. Li, "Cnn-based los/nlos identification in 3-d massive mimo systems," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2491–2494, 2018.
[6] Z. Cui, Y. Gao, J. Hu, S. Tian, and J. Cheng, "Los/nlos identification for indoor uwb positioning based on morlet wavelet transform and convolutional neural networks," *IEEE Communications Letters*, vol. 25, no. 3, pp. 879–882, 2020.
[7] X. Yang, "Nlos mitigation for uwb localization based on sparse pseudo-input gaussian process," *IEEE Sensors Journal*, vol. 18, no. 10, pp. 4311–4316, 2018.
[8] F. Che, Q. Z. Ahmed, J. Fontaine, B. Van Herbruggen, A. Shahid, E. De Poorter, and P. I. Lazaridis, "Feature-based generalized gaussian distribution method for nlos detection in ultra-wideband (uwb) indoor positioning system," *IEEE Sensors Journal*, 2022.
[9] R. Bao and Z. Yang, "Cnn-based regional people counting algorithm exploiting multi-scale range-time maps with an ir-uwb radar," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13704–13713, 2021.
[10] M. J. Kuhn, J. Turnmire, M. R. Mahfouz, and A. E. Fathy, "Adaptive leading-edge detection in uwb indoor localization," in *2010 IEEE Radio and Wireless Symposium (RWS)*, pp. 268–271, IEEE, 2010.
[11] K. Yu, K. Wen, Y. Li, S. Zhang, and K. Zhang, "A novel nlos mitigation algorithm for uwb localization in harsh indoor environments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 686–699, 2018.