eXLoc: Understanding Deep Learning-driven Indoor Localization with eXplainable AI

HongKyeong Jung*, JinYi Yoon†, and HyungJune Lee*
*Department of Computer Science and Engineering, Ewha Womans University, Seoul, South Korea
†Department of Computer Science, Virginia Tech, Blacksburg, VA, USA
Email: hyungjune.lee@ewha.ac.kr

GLOBECOM 2024 - 2024 IEEE Global Communications Conference | 979-8-3503-5125-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/GLOBECOM52923.2024.10901568

***Abstract*— Indoor localization using deep learning has emerged as a promising approach due to its high accuracy in mapping and predicting user locations for complex datasets. However, the inherent complexity of deep learning models often limits their interpretability, creating a gap in user trust and understanding. This paper introduces *eXLoc*, a novel framework that integrates Explainable AI, Class Activation Mapping (CAM), into deep learning models for indoor localization to enhance model transparency and interpretability. We introduce a new metric called *Impact Score* to identify significant APs that affect model predictions. This enhances model interpretability and allows a model to identify the influential APs via their impact on localization performance. We have extensively evaluated *eXLoc* over eight different places from two real-world RSSI datasets. We gained insights into how the model generates predictions, and identified the reasons for the model's poor performance. These results demonstrate that our approach can be effectively utilized in enabling users to have more trust and understanding of the model in many real-world scenarios.**

I. Introduction

With the ubiquitous usage of mobile, edge, and sensor devices, users can easily access various services and information through wireless devices almost anytime and anywhere. This widespread adoption of mobile devices has also had a significant impact on the utilization of location-based services, offering a variety of benefits of enhancing user experiences to facilitating emergency services. Location-based services rely heavily on GPS technology, allowing for real-time tracking of users's current location. However, while GPS excels in providing accurate location data outdoors, it faces a critical challenge in indoor environments. Some factors like signal blockage or reflection within buildings and underground areas often hamper precise location tracking.

To overcome this limitation, indoor localization can be achieved using Wi-Fi networks accessible in various indoor locations such as cafes, museums, and large stadiums. Also, almost all modern devices, such as smartphones and environmental sensors, become equipped with built-in Wi-Fi functionality. For these reasons, WiFi-based localization is widely utilized. Among various methods available for WiFi-based localization, we focus on the fingerprint map that utilizes WiFi signal strength data collected at reference points for localization. The user's location is determined by matching the collected fingerprints with the closest ones at points.

To accurately track the user's location, it is essential to utilize finely collected fingerprint map from various locations. Leveraging this extensive dataset for effective learning is well-suited to deep learning. There have been many research efforts ongoing to improve the accuracy of indoor localization using deep learning techniques such as CNN [1]–[4], LSTM [5], autoencoders [6], [7].

Nevertheless, the complexity of deep learning methods utilized in indoor localization may hinder their interpretability, highlighting the necessity for Explainable AI (XAI) to enhance understanding and trust in these technologies. One example of XAI is Class Activation Mapping (CAM) [8], which visualizes the regions of an image that significantly contribute to the prediction made by a CNN. LIME [9] constructs a simple and interpretable model near the decision boundary of a complex model. This simple model, created through observing the effects of slight changes, is then used to explain the model's prediction for a particular observation or data point. The SHAP [10] approach aims to understand the importance of each feature by considering different combinations of features and measuring the average change in prediction based on the presence or absence of a feature, deriving the Shapley Value.

Applying these XAI to indoor localization allows users to comprehend and trust the model's predictions. Indeed, there have been recent works based on these concepts. For example, recent studies [11], [12] utilize deep learning to enhance the performance of localization and explain the results of localization using XAI techniques such as LIME and SHAP. They use XAI to interpret which APs had an impact on the results. However, they fail to interpret whether some importantly identified APs are specifically correlated with the actual performance of the model. Furthermore, they do not well explain whether these identified important features are actually meaningful or not.

In this paper, we propose *eXLoc*, a framework to understand deep learning-driven indoor localization with eXplainable AI. We leverage CAM [8] to determine the importance of each

This work was supported by the Institute of Information & communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2022-00155966, Artificial Intelligence Convergence Innovation Human Resources Development (Ewha Womans University) and No. RS-2021-II212068, Artificial Intelligence Innovation Hub). The corresponding author is HyungJune Lee.

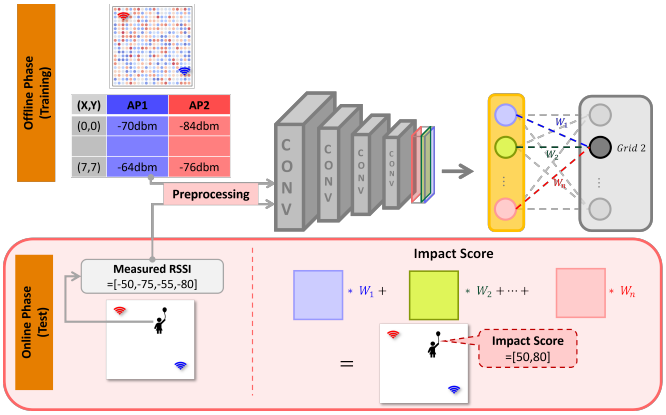


Fig. 1. Our proposed *eXLoc* framework that computes the *Impact Score* across access points to look into the performance of indoor localization accuracy

AP for a certain localization fingerprint map. To combine the CAM with indoor localization, we propose a new metric called *Impact Score*. This score aims to identify a set of influential APs in a fingerprint map. Additionally, we aim to understand how a model makes specific decisions based on specific inputs.

The main contributions of this work are as follows:

- We propose *eXLoc* to explore the actual interpretation of how key APs contribute to localization performance by incorporating the eXplainable AI into indoor localization.
- We introduce the *Impact Score* to identify a set of influential APs that have contributed to prediction. It not only enhances the interpretability of the model by allowing users to understand how decisions are made, but also establishes the validity of interpretability.
- We have evaluated *eXLoc* in eight different places and have demonstrated the interpretability of *Impact Score* in localization. Our *Impact Score* shows the difference between correct and incorrect predictions with a factor of up to 5.11, showing its superior effectiveness in enhancing interpretability compared to 0.94 with LIME.

II. System Overview

We address the problem of understanding deep learning-based indoor localization with eXplainable AI. This work aims to delve into the interpretability of deep learning models for indoor localization. In particular, we take an approach of classifying the correct grid positions in indoor localization, by sub-dividing the region of interests (RoIs) into grids, ensuring robust noise handling and lightweight user localization across extensive areas. In the initial offline phase, we collect RSSI signal data from accessible APs to create a fingerprint for each reference grid point. These fingerprints are then used to construct a fingerprint map, which is the set of fingerprints labeled with the corresponding grid. This map is used to train a CNN, enabling to capture geographical semantics within the grid cells. In the subsequent online phase, the trained model estimates the user location at the grid level by measuring real-time RSSI signals and finding the best-matching RSSI fingerprint.

We propose *eXLoc*, a framework that leverages the eXplainable AI approach for indoor localization with the grid-level classification, as shown in Fig. 1. Without any modification in the model architecture, we localize crucial signal representations in the input RSSI vector by interpreting the resulting trained weights of CNNs. By borrowing an idea from CAM [8], which highlights the impactful pixels in an image, we introduce a new metric called *Impact Score* via global average pooling. We calculate the Impact Score by multiplying each of the feature maps from the last convolutional layer by the weights mapped to a target grid. The Impact Score effectively signifies the relative importance of an AP among other APs in determining a user's location and provides insights behind its prediction. By analyzing the Impact Score, *eXLoc* enables comprehensive interpretation of indoor localization, offering insights into which APs are or should be influential or not, and further contributing to a deeper understanding of CNN-based indoor localization frameworks.

III. Explainable Indoor Localization

Understanding the connection between deep learning and localization has been a challenging problem. To address the problem of cause and effect in deep learning-based localization, we leverage a technique called *CAM* (Class Activation Mapping) [8], which is one of the most effective XAI methods for understanding the impacting pixels in image classification. With the help of CAM framework, we aim to explain the results of localization and investigate why localization accuracy is high at certain areas and low at other areas.

A. Background on CAM

The CNN is a deep learning network commonly used in the field of image processing. Without additional supervised learning, these networks can distinguish specific locations, enabling the extraction of location information within images that leads CNN to predict a particular class. Typically structured with input layers, convolutional layers, and fully connected layers, CNNs excel in classifying images into distinct classes through the utilization of multiple layers of fully connected layers and the softmax function. However, as the features extracted through convolutional layers pass through multiple layers of fully connected layers, the spatial information gets lost, making it challenging to extract object localization information. Therefore, even with high classification accuracy, it is difficult to understand why CNNs have classified specific classes as a result.

The CAM network utilizes Global Average Pooling (GAP) instead of flattening convolutional layers into fully connected layers. This approach generates heatmaps of specific class images, providing insights into how CNNs predict those images into specific classes. The GAP averages all features within each channel, resulting in a vector with as many elements as the number of channels. Subsequently, in applying fully connected layers, the weights obtained are used to create

a heatmap. Each weight is multiplied with each corresponding feature map of the final layer, and the sum yields the heatmap, indicating coordinates in the image that influence the class.

In this paper, we incorporate the CAM into deep learning-based localization by defining and utilizing the *Impact Score*, thereby enhancing the interpretability of the model.

B. Explainable Indoor Localization with CAM

We propose a novel approach for obtaining explainable localization through the integration of CAM. We train the CNN model using fingerprint maps. The fingerprint map represents the RSSI signal patterns for each reference point and is structured as follows.

$$F = \{(r_i, p_i)\}_{i=1}^{M} \tag{1}$$

- p_i: position of a reference point
- r_i: RSSI signal vector measured at the position of p_i
- F: fingerprint map
- M: total number of reference points

When training fingerprints using 1D RSSI vectors with CNNs, there is a problem of losing edge information within the data. To address this problem, we create cyclic 2D radio maps by cyclically reshaping the given 1D input data. This approach allows the model to consider all of the information in the input data and can enhance performance by substituting for lost information. The process of creating a cyclic 2D radio map can be described as follows:

$$\text{Cyclic 2D map}[R, C] = \text{RSSI}[(C - R) \bmod N] \tag{2}$$

where R represents the row index, C represents the column index, where $R \geq 0$ and $C \geq 0$. N denotes the number of APs. This radio map is derived from the given 1D fingerprint, where each column represents the values of the original fingerprint. Each row of the radio map displays the result of cyclically shifting the original fingerprint.

The CNN model is designed to classify the user's location based on cyclic 2D radio maps. In this process, the CNN model learns to extract spatial features from the input cyclic 2D map and associate them with the corresponding user locations. Then, it adjusts its parameters to minimize the prediction error between the predicted location and the ground truth location. Once trained, the CNN model can effectively predict the user's location based on the given cyclic 2D maps.

When the model makes a prediction to be class c, we analyze which APs have significantly contributed to the outcome c via the "window" of the Impact Score. The procedure for calculating the Impact Score S_c of class c is described below.

First, after predicting the user's location, the average value of each feature map $F_k(x, y)$ in the last convolutional layer is calculated. Here, k represents the index of the feature map, while x and y denote the spatial coordinates within the feature map. This is performed using the following equation:

$$GAP(F_k) = \frac{1}{H \times W} \sum_{x=1}^{H} \sum_{y=1}^{W} F_k(x, y). \tag{3}$$

TABLE I
SUMMARY OF DATASETS FOR INDOOR LOCALIZATION

Place	RoI Size (m^2)	# of Training data	# of Test Data	# of APs
Hall	8.1×8.0	17,284	4,322	10
Lounge	6.6×9.9	4,889	1,223	12
Office-A	9.9×9.9	6,828	1,708	16
Office-B	6.3×10.8	6,048	1,512	10
B1-A	3.0×9.0	339	152	11
B1-B	4.0×4.0	348	147	27
1F	7.0×5.0	235	42	11
2F	3.0×11.0	635	312	11

We utilize the weights w_k^c for each F_k, where w^c represents the weights specific to class c. After multiplying each feature map F_k by its corresponding weight w_k^c, we perform a pixel-wise sum to obtain the contribution of each feature map to the prediction for class c. This process combines the spatial information represented by F_k with the importance assigned to each feature map by the weights w_k^c specific to class c. The contribution of each feature map is calculated as:

$$R_c(x, y) = \sum_{k,y} \sum_{k} w_k^c \cdot F_k(x, y). \tag{4}$$

Subsequently, we resize the result to match the input size, denoted as $R_{resized}$. Now, we compute the Impact Score by averaging values cyclically shifted from each position in $R_{resized}$.

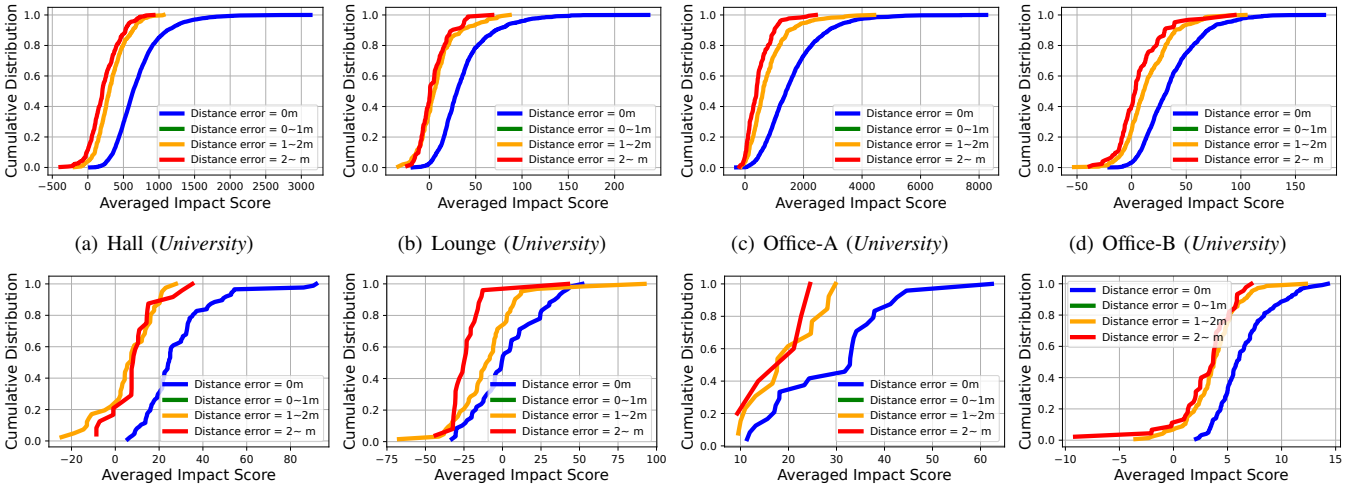
$$\text{Impact Score}_c[i] = \frac{1}{N} \sum_{j=0}^{N-1} R_{\text{resized}}[j,\ (i + j) \bmod N]. \tag{5}$$

Here, Impact Score$_c[i]$ represents the value at the i th position in the resized 1D array indicating the impact score value of AP_i and N represents the number of APs. Therefore, Impact Score$_c[i]$ represents the influence score of class c at a specific AP_i. This process enables the interpretation of localization results based on the CNN-based model without altering its network structure. The Impact Score highlights influential areas for predicting each class, providing valuable insights into CNN's decision-making process. In other words, when determining the user's location, the Impact Score offers an effective way to assess of which APs are influential or not for a certain area.

IV. EXPERIMENTS

We validated *eXLoc* on RSSI data of Wi-Fi signals collected in eight places from two different datasets: 1) *University*: Hall, Lounge, Office-A, and Office-B; and 2) *Microsoft Indoor* [13]: B1-A, B1-B, 1F, and 2F. The details are explained in Table I.

To enhance robustness against outliers and to reduce the impact of Wi-Fi signal fluctuations, we preprocess the RSSI values by normalizing into z-score for each place [14], [15]. Additionally, we rearrange the input vectors particularly for University datasets using M-TSP (Multiple Traveling Salesman Problem) [16], which is a heuristic approach to minimize the travel route for visiting all APs.



(a) Hall (*University*) (b) Lounge (*University*) (c) Office-A (*University*) (d) Office-B (*University*)

(e) B1-A (*Microsoft Indoor*) (f) B1-B (*Microsoft Indoor*) (g) 1F (*Microsoft Indoor*) (h) 2F (*Microsoft Indoor*)

Fig. 2. The cumulative distribution of the average of three highest Impact Score for each distance error interval

TABLE II
LOCALIZATION PERFORMANCE OF REFERRED CNN ARCHITECTURE

Place	Accuracy (%)	Loose Acc (%)	Distance Error (m)
Hall	82.1	90.4	0.46
Lounge	81.6	90.1	0.40
Office-A	73.2	86.5	0.58
Office-B	70.4	88.2	0.53
B1-A	57.2	75.7	0.82
B1-B	40.1	63.3	0.93
1F	57.1	78.6	0.69
2F	39.1	59.0	1.36

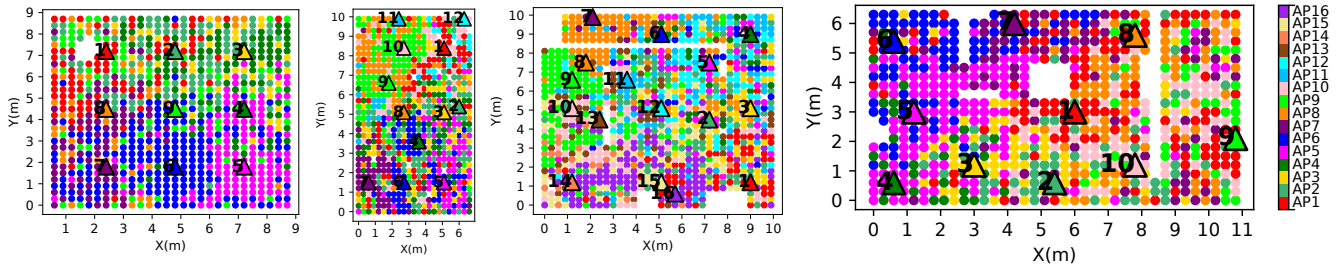
CNN architecture. We implemented *eXLoc* on a CNN [17]-based indoor localization consisting of 6 hidden layers of 512 neurons with ReLU activation function and Adam optimizer, with a learning rate of 0.001. For RSSI vectors as inputs, we transform the vectors with the length of N into cyclic 2D radio maps of $N \times N$ values based on Eq. (2). We sub-divide RoI into grids with the size of $1 \times 1\ m^2$, and the model expects the output class of the most probable grid. For example, the RoI with the size of $8.1 \times 6.3\ m^2$ consists of $\lceil 8.1 \rceil \times \lceil 6.3 \rceil = 63$ classes.

Evaluation metrics. We used three metrics to evaluate the localization performance: 1) Accuracy (%): the accuracy of prediction; 2) Loose Acc (%): the accuracy of predictions within a tolerance of 3×3 grids; and 3) Distance Error (m): the distance between the centroids of the predicted grid and the ground-truth grid. The localization performance of used CNN-based indoor localization is in Table II.

A. Analysis on Impact Score

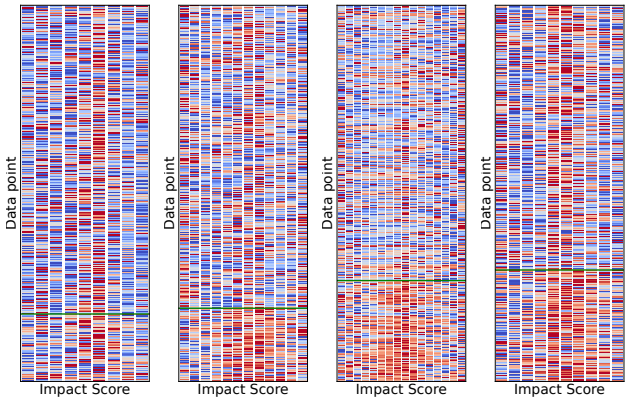
The correlation between Impact Score and the localization performance. We showed the cumulative distribution of the averaged Impact Scores with respect to the distance error in Fig. 2. We categorize the data points into the following groups: 1) correct answers, where the distance error is zero; 2) points with distance error less than 1 m; 3) points with distance error less than 2 m and greater than or equal to 1 m; and 4) points with distance error greater than or equal to 2 m. Considering the trilateration for localization, we computed the average of the three highest Impact Scores, which have the most significant impact on localization. There were no data points with a distance error in the range of 1 to 2 meters, as indicated as the green line. As shown in Figs. 2(a)-2(d) within *University* and Figs. 2(e)-2(h) within *Microsoft Indoor*, the accurately predicted group consistently has a higher impact score, while the Impact Scores decrease for the other groups as the distance error increases. This highlights the importance of the Impact Score in finding the effectiveness of localization, implying that the Impact Score can be utilized as a key metric for evaluating positioning systems. It also means that even without ground-truth data to compute the distance error, $eXLoc$ offers a way to estimate the relative quality of localization through this Impact Score.

AP-wise Impact Score. We explored the distribution of selected APs, which are the closest among the top three APs with the highest Impact Scores for each location in Fig. 3. In general, users are located closer to AP, resulting in a stronger signal. In this context, we closely examined how the APs with the high Impact Score are distributed regarding their distance from the APs, which correlates with the signal strength. As shown in Fig. 3(a), the data points are quite clearly clustered with the same color. In particular, for AP #5 in magenta, AP #6 in blue, and AP #7 in purple, the data points are more likely to have higher Impact Scores when they are closer to APs with stronger signal strength. Also, the distinction is clearer in Hall, where localization performance is higher. It implies a potential correlation among the signal strength, the Impact Score, and localization fidelity. However, as we pointed out in Fig. 2, localization requires at least three signal values to decide a certain location. This means that justifying the Impact Score with a single AP may be insufficient. Furthermore, the single



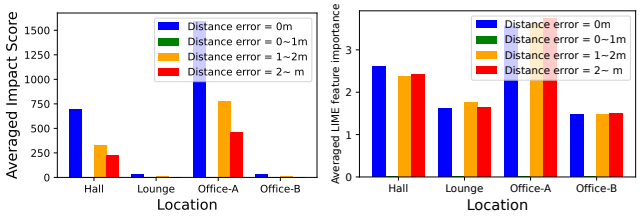
(a) Hall (b) Lounge (c) Office-A (d) Office-B

Fig. 3. Distribution map of the closest AP among the top three APs with highest Impact Scores for places in *University* dataset, where the AP locations are represented by triangles with corresponding IDs



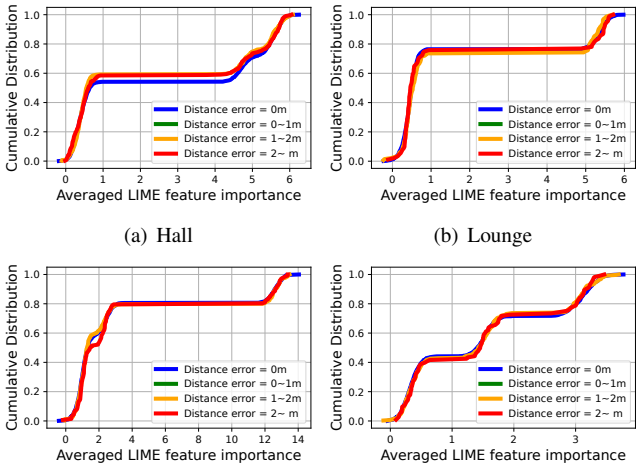
(a) Hall (b) Lounge (c) Office-A (d) Office-B

Fig. 4. Distribution of the Impact Scores of APs for places in *University* dataset, where the red indicates the higher Impact Score, while the blue indicates the lower Impact Score. The data points above the green line represent correct predictions, and they are sorted in ascending order based on the distance error



(a) CAM (b) LIME

Fig. 5. The comparison of the averaged values of the three highest Impact Scores between CAM and LIME



(a) Hall (b) Lounge

(c) Office-A (d) Office-B

Fig. 6. The cumulative distribution of the average of three highest Impact Scores for each distance error interval using LIME [9], in contrast to our *eXLoc* (as in Fig. 2)

signal value has a risk of fluctuation due to the obstacles, as illustrated in Figs. 3(b)-3(d). With more obstacles from Hall to Office-B, there are more possibilities to have noisy signal values, leading to less distinct clusters among APs. This suggests that there is a rough correlation among distance from APs, signal clarity/stability (rather than strength), and the Impact Score. While we observe a correlation for each AP perspective in Fig. 3, the cooperation of a subset of at least three APs demonstrates a synergistic correlation with localization accuracy, as illustrated in Fig. 2.

Analysis on incorrect prediction. We investigated the Impact Score of APs for each data point, sorted by distance error, in Fig. 4 to explore the variations in Impact Scores between the correct and the incorrect predictions. As suggested in CAM, pinpointing specific pixels of interest is important for accurate predictions. However, different from the correct predictions above the green line, incorrect predictions exhibit numerous signal values with high Impact Scores spread in red. It means that the referenced CNN model may struggle to identify the key input signal values to focus on, leading to confusion with almost similar high impact scores for prediction. This difference between correct and incorrect predictions in terms of Impact Scores highlights the importance of accurately localizing key signal features for successful predictions. This suggests explainable AI techniques in identifying specific pixels are essential for accurate localization. Also, simply having a high Impact Score does not necessarily ensure better predictions. Instead, it is critical to pinpoint the important subsets of inputs that are truly relevant and find somewhat high deviations among APs in terms of impact score.

Comparison to other explainable AI algorithms. We employed a different explainable AI approach, Local Inter-

pretable Model-agnostic Explanations (LIME) [9], [11], [12], to assess the effectiveness of our CAM-based framework in understanding deep learning-based localization. We presented the averaged Impact Scores of the top three APs for LIME and CAM in Fig. 5 and the cumulative distribution for each distance error group using LIME in Fig. 6. LIME simplifies the interpretation of a complicated black-box model by constructing a lightweight explaining model such as linear regression, and the coefficients of this formulated interpretable model indicate the importance of each feature. We implemented LIME using the same input format as CAM, which consists of 1D RSSI vectors. As depicted in Fig. 5(a), CAM has clear high Impact Scores for the correct prediction group, whereas the wrong prediction groups exhibit lower Impact Scores. However, as illustrated in Fig. 5(b), LIME does not show a distinct difference across different distance errors. To take a closer look at the distribution, LIME in Fig. 6 exhibits a very similar distribution across different distance errors, in contrast to CAM in Fig. 2. This clear distinction in Impact Scores in our CAM-based $eXLoc$ suggests its effectiveness in interpreting deep learning-based localization models.

B. Discussion

Our explainable AI approach based on CAM for deep learning-based localization proves effectiveness in understanding the correlation of the synergic Impact Scores with the distance error rather than the single signal value itself, as shown in Fig. 2 and Fig. 3. This capability is particularly valuable in applications where location awareness is important, such as indoor navigation and smart homes. Furthermore, the Impact Score shows its ability to evaluate the quality of each signal value from APs, as shown in Fig. 4. It can contribute to enhancing network performance and improving user experience. For example, identifying APs with low Impact Scores enables us to enhance coverage in specific areas by installing additional APs or optimizing the placement of existing APs.

Lastly, from the insights from Fig. 5 and Fig. 6, CAM verifies its effectiveness in interpreting deep learning-based localization models. However, given the diversity of explainable AI techniques available, these observations stress the necessity for techniques capable of pinpointing important parts in input data correlated to the distance error. This suggests that deep learning-based systems may benefit from employing suitable XAI techniques to enhance interpretability and understanding.

V. Conclusion

We propose *eXLoc*, a framework to understand deep learning-based indoor localization by introducing Impact Score by CAM to identify the significant APs in the model predictions. Our *eXLoc* improves the interpretability of the deep learning model, enabling users to understand the decision-making process. Based on analysis on eight different real-world RSSI datasets, we have validated that *eXLoc* demonstrates the correlation between Impact Scores and localization performance. Additionally, we disclose the unrevealed reasons contributing to the incorrect predictions.

For future work, we can further leverage Impact Scores to not only understand localization but also enhance the prediction. It would be interesting to explore localizing the user using the robust signal data by Impact Scores or investigate the optimal AP placement strategies. Furthermore, exploring dynamic and large-scale networks could enhance robustness and extend applicability to real-world scenarios.

References

[1] J.-W. Jang and S.-N. Hong, "Indoor localization with wifi fingerprinting using convolutional neural network," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2018, pp. 753–758.

[2] W. Qian, F. Lauri, and F. Gechter, "Supervised and semi-supervised deep probabilistic models for indoor positioning problems," *Neurocomputing*, vol. 435, pp. 228–238, 2021.

[3] X. Song, X. Fan, C. Xiang, Q. Ye, L. Liu, Z. Wang, X. He, N. Yang, and G. Fang, "A novel convolutional neural network based indoor localization framework with wifi fingerprinting," *IEEE access*, vol. 7, pp. 110 698–110 709, 2019.

[4] W. Njima, I. Ahriz, R. Zayani, M. Terre, and R. Bouallegue, "Deep cnn for indoor localization in iot-sensor systems," *Sensors*, vol. 19, no. 14, p. 3127, 2019.

[5] Z. Chen, H. Zou, J. Yang, H. Jiang, and L. Xie, "Wifi fingerprinting indoor localization using local feature-based deep lstm," *IEEE Systems Journal*, vol. 14, no. 2, pp. 3001–3010, 2019.

[6] Z. E. Khatab, A. Hajihoseini, and S. A. Ghorashi, "A fingerprint method for indoor localization using autoencoder based deep extreme learning machine," *IEEE sensors letters*, vol. 2, no. 1, pp. 1–4, 2017.

[7] K. S. Kim, S. Lee, and K. Huang, "A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on wi-fi fingerprinting," *Big Data Analytics*, vol. 3, pp. 1–17, 2018.

[8] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.

[9] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[10] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[11] A. G. Kakisim, Z. Turgut, and T. Atmaca, "Xai empowered dual band wi-fi based indoor localization via ensemble learning," in *2023 14th International Conference on Network of the Future (NoF)*. IEEE, 2023, pp. 150–158.

[12] Z. Turgut and A. G. Kakisim, "An explainable hybrid deep learning architecture for wifi-based indoor localization in internet of things environment," *Future Generation Computer Systems*, vol. 151, pp. 196–213, 2024.

[13] Y. Shu, Q. Xu, J. Liu, R. R. Choudhury, N. Trigoni, and V. Bahl, "Indoor location competition 2.0 dataset," January 2021. [Online]. Available: https://www.microsoft.com/en-us/research/publication/indoor-location-competition-2-0-dataset/

[14] M. Nowicki and J. Wietrzykowski, "Low-effort place recognition with wifi fingerprints using deep learning," in *Automation 2017: Innovations in Automation, Robotics and Measurement Techniques 1*. Springer, 2017, pp. 575–584.

[15] Y. Lin, K. Yu, L. Hao, J. Wang, and J. Bu, "An indoor wi-fi localization algorithm using ranging model constructed with transformed rssi and bp neural network," *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 2163–2177, 2022.

[16] T. Bektas, "The multiple traveling salesman problem: an overview of formulations and solution procedures," *Omega*, vol. 34, no. 3, pp. 209–219, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305048304001550

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.