Efficient and Accurate Indoor Positioning System: A Hybrid Approach Integrating PCA, WKNN, and Linear Regression

Thi Hang Duong[1, 2], Anh Vu Trinh[2], and Manh Kha Hoang[1, *]

[1] Faculty of Electronics and Engineering, Hanoi University of Industry, Ha Noi, Viet Nam
[2] Department of Electronics and Telecommunication, VNU University of Engineering and Technology (VNU-UET), Ha Noi, Viet Nam
Email: hangdt@haui.edu.vn (T.H.D.); anhvutrinh1811@gmail.com (A.V.T.); khahoang@haui.edu.vn (M.H.H.)
*Corresponding author

***Abstract*—The high-precision Indoor Positioning System (IPS) is a captivating area of research that has made significant advancements in recent years due to the increasing demand for its applications. Our study proposes an innovative approach to improve indoor positioning accuracy by integrating Principal Component Analysis (PCA), weighted k-nearest Neighbors (WKNN), and Linear Regression (PCA-WLR). This hybrid strategy enables the system to leverage the unique characteristics of each model, capturing intricate patterns and correlations in the data. Experimental evaluations on a publicly available dataset demonstrate the superiority of our hybrid approach. The Root Mean Squared Error (RMSE) achieved is 1.97 meters, and the mean distance error is 2.23 meters. Remarkably, the ensemble outperforms individual methods in other studies on the same dataset, showing 10.8% to 17.2% improvement in accuracy. Notably, our proposed hybrid approach significantly reduces training time from 581.3599 seconds to 8.8814 seconds, representing an impressive reduction of approximately 98.47%. Similarly, testing time is reduced from 10.1721 seconds to 0.0176 seconds, indicating a substantial decrease of around 99.82%. These significant reductions in training and testing times underscore the efficiency and effectiveness of our proposed ensemble model, making it highly practical for real-time applications.**

***Keywords*—indoor localization, Principal Component Analysis (PCA), Weighted k-nearest Neighbors (WKNN), linear regression, ensemble model, reduced-dimensional**

I. INTRODUCTION

Indoor positioning has garnered substantial attention in recent years due to its wide-ranging applications in navigation, tracking, and location-based services. The ability to accurately estimate the position of devices within indoor environments is crucial for enabling various services and enhancing user experiences [1, 2]. However, indoor positioning remains a complex assignment, primarily due to the complexities associated with indoor environments, such as signal interference, multipath propagation, and limited line of sight.

To address these challenges, numerous approaches have been proposed in the field of indoor positioning. One promising avenue is the use of machine learning techniques, which have shown great potential in improving the accuracy and robustness of indoor positioning systems [2, 3]. Numerous machine learning and deep learning techniques have been introduced to enhance the accuracy and effectiveness of indoor localization, including CNN, LSTM, k-NN, WKNN... Among these methodologies, Convolutional Neural Networks (CNNs) have garnered significant attention for their adeptness in automatically learning spatial features from raw data [4, 5]. Conversely, Long Short-Term Memory (LSTM) networks have demonstrated their effectiveness in capturing temporal dependencies in sequential data, rendering them ideal for time-series localization tasks [6]. K-Nearest Neighbors (KNN) and Weighted K-Nearest Neighbors (WKNN) [7–9] stand as non-parametric algorithms commonly employed in localization, owing to their simplicity and proficiency in handling noisy and non-linear data.

Among these techniques, PCA, Weighted k-Nearest Neighbors (WKNN) [9], and Linear Regression have been widely explored for their effectiveness in handling high-dimensional data, capturing spatial patterns, and establishing relationships between input features and target outputs.

In this study, we propose an ensemble-based approach that combines the strengths of PCA, WKNN, and Linear Regression to enhance indoor positioning estimation. The ensemble framework allows us to harness the complementary aspects of these individual models, thereby creating a more powerful and accurate positioning system. Specifically, PCA is utilized to reduce the dimensionality of the input data, mitigating the curse of dimensionality and eliminating redundant or irrelevant features. The reduced-dimensional feature space is then fed into the ensemble model, comprising WKNN and Linear Regression, to capitalize on their respective capabilities in handling proximity-based localization and modeling linear relationships.

Manuscript received July 26, 2023; revised August 16, 2023; accepted September 13, 2023; published January 23, 2024.

The key contributions of our work are as follows:

Improved Positioning Accuracy: Experimental results on a publicly available dataset demonstrate the superiority of our hybrid approach. We achieved a Root Mean Squared Error (RMSE) of 1.97 meters and a mean distance error of 2.23 meters. Notably, our hybrid approach outperforms contemporary methods [6] on the same dataset, showcasing an accuracy improvement ranging from 10.8% to 17.2%.

Reduced Computational Time: Our hybrid solution significantly reduces training time from 581.3599 seconds to 8.8814 seconds, representing a remarkable reduction of over 98.47%. Similarly, testing time is reduced from 10.1721 seconds to 0.0176 seconds, indicating a substantial decrease of approximately 99.82%. These substantial reductions in training and testing times underscore the efficiency and practicality of our proposed ensemble model, making it highly suitable for real-time applications.

The remainder of this paper is organized as follows: Section II presents a comprehensive review of related work in indoor positioning and machine learning techniques. Section III details the materials and methods. Section IV presents the experimental setup and discusses the obtained results. Finally, Section V concludes the paper and outlines potential avenues for future research in the field of indoor positioning.

II. Literature Review

The field of high-precision Indoor Positioning Systems (IPS) has witnessed remarkable progress in recent years due to its increasing significance and diverse applications. IPS plays a pivotal role in various applications, such as asset tracking, navigation assistance, and location-based services. However, traditional IPS methods often encounter challenges in dealing with the complexities of indoor environments, including signal interference, multipath propagation, and limited line of sight. To address these challenges and improve indoor positioning accuracy, researchers have turned to innovative approaches that leverage the capabilities of machine learning and deep learning techniques.

In recent years, several machine learning models, such as Convolutional Neural Networks (CNN) [1, 5, 10], Long Short-Term Memory (LSTM) [6, 11, 12] networks, k-Nearest Neighbors (KNN), and weighted k-Nearest Neighbors (WKNN) [8, 9, 13], have been proposed to enhance indoor localization accuracy. Dai *et al.* [14] used the DNN to divide the area into four subareas, and an improved K-nearest neighbor (KNN) is used to determine the location. In a previous study [5], a Wi-Fi fingerprinting method based on Convolutional Neural Networks (CNN) demonstrated superior performance compared to Deep Neural Network (DNN)-based methods. Nabati *et al.* [15] proposed a deep learning model, called generative adversarial networks, to learn the distribution of limited data in a four-class problem. The learned DNN model is then used to generate synthetic data, which can be combined with real data to enhance overall localization accuracy. Song *et al.* [16] introduced a groundbreaking approach by integrating Convolutional Neural Network (CNN) and Stacked Auto-Encoder (SAE) to tackle the indoor localization task using Wi-Fi signals. The authors proposed combining CNN and SAE to achieve enhanced precision and efficiency in indoor localization across multi-building and multi-floor environments. By leveraging the strengths of both CNN and SAE, their proposed method offers significant improvements in the accuracy and efficiency of indoor positioning. Chen *et al.* [12] introduced a local feature-based deep long short-term memory (LF-DLSTM) approach for Wi-Fi fingerprinting indoor localization, focusing on local feature extraction to mitigate noise effects and obtain robust features. Wang *et al.* [8] introduced a novel WKNN method that addresses the RSS similarity and position distance relationship. It calculates the weighted Euclidean distance (WED) considering the RSS and signal propagation distance differences based on the spatial signal attenuation law. The approach then derives the approximate position distance (APD) using the position distances and WEDs between Reference Points (RPs). In [6], a classifier scheme for RSS-based indoor positioning using RNN is introduced, along with the implementation of LSTM as a variant of RNN. Zheng *et al.*[17] proposed a self-calibration deep learning framework using auto-encoders to reduce the effect of environmental changes on localization performance. This approach demonstrates significant improvements in the localization problem on a publicly available dataset. Since we share the same dataset with this study, all of its results are chosen for comparison with our proposed approach. The comparisons encompass positioning accuracy as well as training and testing times.

Generally, various machine learning models, such as CNN, LSTM, KNN, and WKNN, have been proposed to improve indoor localization accuracy. While CNN has shown superior performance in some cases, its computational complexity can be a limitation, leading to longer training times and inference latency, making real-time positioning challenging. Additionally, local feature-based approaches, like LF-DLSTM, focus on extracting robust features to mitigate noise effects. Many DNN approaches primarily focus on learning the distribution of RSS samples within subareas, rather than at each Reference Point (RP) individually during the offline phase of positioning.

To address these limitations and improve indoor positioning, this paper proposes a novel approach that combines PCA, WKNN, and Linear Regression. By leveraging the unique strengths of each model, this hybrid strategy aims to achieve enhanced precision and efficiency in indoor localization across multi-building and multi-floor environments. The PCA helps in capturing intricate patterns and correlations in the data, while WKNN enables more accurate and efficient selection of the nearest reference points. Finally, Linear Regression is utilized for precise estimation of the user's position.

III. Materials and Methods

In this section, we present a concise overview of the dataset and the principle of data collection. Subsequently, we showcase the PCA algorithm utilized to extract the local features. After that, we introduce the PCA-WLR architecture, specifically designed for indoor location determination through Wi-Fi fingerprinting. Lastly, we delve into the training process of the PCA-WLR approach and address relevant aspects.

A. Dataset

The dataset was collected in a library with expansive bookshelves on the 3rd and 5th floors, covering an area exceeding 300 m^2 [18]. To ensure data accuracy and address potential issues resulting from signal variations over time, data collection extended over 15 months, encompassing more than 60,000 measurements. Each entry in the dataset includes crucial information, such as the object's location, RSSI indicators from Wi-Fi AP access points, execution time, and identification data. Throughout the measurements, 448 Wi-Fi AP access points were detected, situated on both the 3rd and 5th floors, approximately 2.65 meters above the ground.

To generate a wireless positioning map, various locations were meticulously chosen, and fingerprint data was captured from four different orientations: front, back, left, and right. During the offline training phase, the data collector stood at a known reference point while holding a Samsung Galaxy S3 phone in their right hand, positioned in front of their chest. The mobile phone was equipped with an application designed to collect Wi-Fi RSSI data, streamlining the data collection process.

The dataset was subsequently partitioned into two sets for training and testing purposes. The training dataset consisted of 16,704 fingerprints obtained from 24 distinct reference points, while the test dataset encompassed 46,800 fingerprints collected from 106 different reference points. Each fingerprint in the dataset consisted of 448 RSSI indicators from the corresponding access point.

B. Feature Extraction using PCA

To tackle the challenges posed by high-dimensional Wi-Fi localization datasets, the implementation of dimensionality reduction techniques, such as PCA or Autoencoder, has proven to be an effective solution. These methods reduce data complexity and eliminate non-essential features, resulting in a more streamlined and informative data representation. As a consequence, this enhanced data representation not only improves the efficiency and accuracy of the localization process but also addresses issues related to sparse data and missing values. In this section, we introduce the use of PCA in our Wi-Fi-based localization to reduce dimensionality, showcasing their significant contributions to the overall performance of the system. The PCA algorithm is presented in the Algorithm 1.

C. Prososed Approach

1) The block diagram of proposed approach

Algorithm 1: PCA Algorithm

Input:

Assuming X is the data matrix n x m (n=16704, m=448) and p is the desired number of components (principal components).

- Data matrix **X** (with each row representing a data point and each column representing a feature)

- Number of desired components p (p < D, where D is the original data's dimensionality) (D=448, and p =30).

Output:

- Matrix containing the data projected onto the p principal components

Steps**:**

Step 1: Normalize the data by subtracting the mean and dividing by the standard deviation for each feature. Suppose the normalized data matrix is denoted as $\mathbf{X}_{norm}$ with a size of (n x m).

Step 2: Calculate the covariance matrix S of the normalized data using the formula:

$$S = \frac{1}{n-1} X_{norm}^{T} X_{norm}$$

$$\mathbf{X}_{reduced} = \mathbf{X}_{norm}\mathbf{U}$$

where X_{norm}^{T} is the transpose of X_{norm}, n is the number of samples, and S has a size of (D × D**)**.

Step 3: Compute the eigenvectors V and eigenvalues D of the covariance matrix S using the formula:

S × V = V × D

where **V** is the matrix containing the eigenvectors of **S** as its columns, and **D** is the diagonal matrix containing the eigenvalues of **S**. Both **V** and **D** have a size of (DxD).

Step 4: Sort the eigenvalues in descending order and select the top p (p=30) eigenvalues. Choose the corresponding p (p=30) eigenvectors to form the matrix **U** with a size of (Dxp).

Step 5: Reduce the data dimension from D features to K features using the formula:

$$\mathbf{X}_{reduced} = \mathbf{X}_{norm}\mathbf{U}$$

where $\mathbf{X}_{reduced}$ is the data matrix that has been reduced with a size of (nxp)

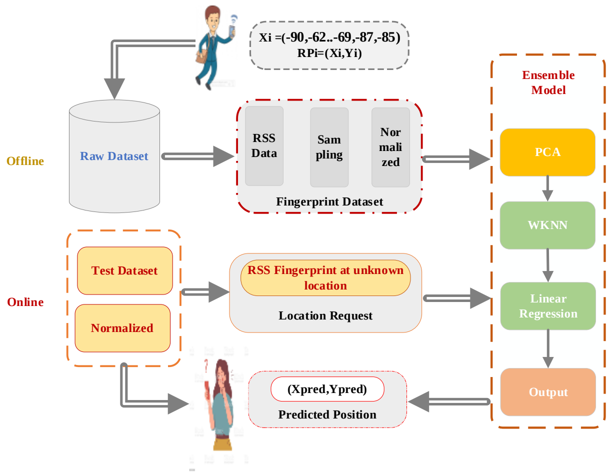

Fig. 1. The block diagram of proposed approach.

Fig. 1 describes the block diagram of the proposed approach. The diagram illustrates a multistep indoor positioning process using the PCA-WLR approach. The steps are as follows:

PCA (Dimensionality Reduction):

- Apply PCA to reduce the data from 448 dimensions to 30 dimensions (p = 30).
- Retain principal components explaining the most variance to reduce complexity while preserving vital information.

WKNN (Localization using k-Nearest Neighbors with Weights):

- Utilize the reduced data (30 dimensions) from PCA as input for WKNN.
- Apply inverse distance weighting to emphasize closer neighbors' importance in estimating the target position.

Linear Regression (Localization Refinement):

- Combine the estimated position from WKNN with the 30-dimensional reduced data from PCA.
- Employ Linear Regression to predict the accurate coordinates of the target by modeling the relationship between the reduced data and the estimated position.

Training and Testing Phases:

- Train the PCA, WKNN, and Linear Regression models using a labeled dataset containing 16704 coordinates and their corresponding RSSI values.
- In the testing phase, apply the trained models to new data from 46800 coordinates with their respective RSSI values, following the same PCA, WKNN, and Linear Regression steps as in the training phase.
- The testing phase provides the final predicted coordinates for indoor positioning.

2) The proposed approach algorithm

In this section, we introduce our proposed ensemble approach for predicting the position of an object. The process involves several steps. Firstly, we apply PCA to reduce the dimensionality of the data, enabling us to capture the essential features efficiently. Next, the WKNN (Weighted k-Nearest Neighbors) algorithm is employed for predicting the object position based on the k-nearest neighbors in the feature space. Subsequently, Linear Regression is used to predict the final coordinates.

Within the WKNN algorithm, it is important to emphasize that closer data points carry higher significance in the classification decision-making process. The parameter k in the WKNN algorithm determines the number of nearest neighbors used to classify a new data point.

When the value of k is 1, the model relies solely on the nearest neighbor for prediction. However, this can lead to "overfitting," wherein the model becomes too sensitive to noise in the data. Consequently, the model may be influenced excessively by outliers, leading to difficulties in generalizing new data. On the other hand, as the value of k becomes larger and approaches the number of data points in the training set, the model becomes less sensitive to noise. Nevertheless, employing a large k value may result in the model overlooking complex boundaries, which can lead to "underfitting" and limited predictive capabilities. Selecting an appropriate value of k is a critical step in effectively deploying the WKNN algorithm. A common approach to determine the optimal k value is to use cross-validation, dividing the data into multiple training and validation sets, and subsequently evaluating the model's performance on these sets. This iterative process allows us to find the k value that strikes the right balance between capturing essential patterns in the data while avoiding overfitting or underfitting issues. By doing so, we can enhance the accuracy and robustness of our predictive model.

Fig. 2 illustrates the relationship between the value of k and the distance error. The results highlight that the minimum distance error occurs when k is 30. The findings indicate that values of k smaller or larger than this optimal value result in higher distance errors. This analysis is of utmost significance, as it provides valuable insights into the impact of the choice of k on the accuracy of our prediction model.

Algorithm 2: The Linear Regression Algorithm

Input: Dataset obtained after performing dimensionality reduction using PCA. The input data vectors have the form:

$$X_{iLR} = \left(x_{i1PCA}, x_{i2PCA} ..., x_{ikPCA}, x_{i1WkNN}, x_{i2WkNN}\right)$$

In which, $\left(x_{i1PCA}, x_{i2PCA} ..., x_{ikPCA}\right)$ is the data vector after dimensionality reduction by PCA, $\left(x_{i1WkNN}, x_{i2WkNN}\right)$ is the position coordinates predicted by the WKNN algorithm at k which is 30.

Output: $Y_{LR} = \left(x_{Pred}, y_{Pred}\right)$

Step 1: Data Preparation:

- Collect data with 32 input features:

$$X_{iLR} = \left(x_{i1PCA}, x_{i2PCA} ..., x_{ikPCA}, x_{i1WkNN}, x_{i2WkNN}\right)$$

- Divide the data into two sets: the training set and the test set. The training set will be used to train the model, while the test set will be used to evaluate the performance of the trained model;

Step 2: Model Initialization

- Initialize the weight vector (w) and bias term (b) with random values or set them to zero.

Step 3: Training the Model:

For each data point in the training set:

- Compute the predicted output x_{Pred}, y_{Pred} using the current weight vector and bias. Compute the predicted position coordinates by Linear Regression Algorithm:

$$\hat{x}_{Pred} = b_{01} + w_i x_{ikPCA} + x_{i1WkNN}$$

$$\hat{y}_{Pred} = b_{02} + w_i x_{ikPCA} + x_{i2WkNN}$$

In which: b_{01}, b_{02} are bias values, $w_1, ..w_k\,(k = 30)$ are the coefficients corresponding to the features after dimensionality reduction using PCA. x_{i1WkNN}, x_{i2WkNN} are the prediction results of the WKNN model for the two outputs X and Y at the respective data points. The Linear Regression model will optimize the coefficients $w_1, ..w_k\,(k = 30)$ based on the training set and the prediction results of the WKNN model to predict the X and Y coordinates of the object on the test data. Calculate the errors between the predicted output and the true output. Update the weight vector and bias using gradient descent to minimize the Mean Squared Error (MSE).

Step 4: Prediction

- After training the model, use the learned weight vector and bias to make predictions for new data points:

$$x_{Pred} = b_{01} + w_i x_{iknew} + x_{iWkNNnew}$$

$$y_{Pred} = b_{02} + w_i x_{iknew} + x_{i2WkNNnew}$$

By examining the distance error across different values of k, we can determine the optimal k value that yields the most accurate predictions. The observed trend of the distance error decreasing at k is 30, and then increasing for values deviating from this optimal value, guides us in selecting an appropriate k value to achieve superior performance. Setting k to 30 is a critical decision that strikes the right balance between overfitting and underfitting issues. It ensures that the model captures relevant patterns in the data while maintaining generalizability to new instances. Thus, using WKNN with k = 30 showcases its effectiveness in accurate position prediction in the first step of our approach.

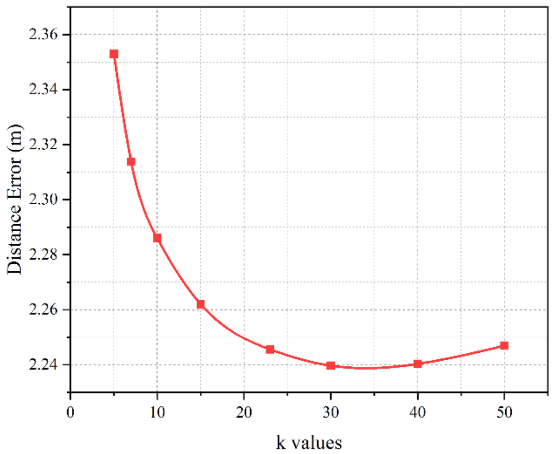

Fig. 2. The relationship between the value of k and the distance error.

We use the Linear Regression algorithm presented in Algorithm 2 to predict the final position.

IV. Result and Discussion

In this section, we presented results and discussion a dimensionality reduction method using PCA combined with a machine learning model consisting of Weighted k-Nearest Neighbors (WKNN) and Linear Regression (PCA-WLR) to predict indoor localization.

We conducted experiments on multiple models, adjusting the number of components (k) in the WKNN algorithm and the dimensionality of data reduction in the PCA algorithm to determine the best-performing model before comparing it with other approaches. The result shows that the PCA-WLR approach compared with the LSTM algorithm [6], the positioning accuracy is significantly improved, and the PCA-WLR algorithm is more robust. Using the mean distance error (MDE) describe algorithm performance, the actual position of the object to be measured by (x_{itrue}, y_{itrue}). Target estimated position is (x_{ipred}, y_{ipred}), N_{test} is the total samples. The MDE calculation is as follows Eq. (1):

$$MDE = \frac{\sum_{i=1}^{N_{test}} \sqrt{\left(x_{ipred} - x_{itrue}\right)^2 + \left(y_{ipred} - y_{itrue}\right)^2}}{N_{test}} \quad (1)$$

The simulation results demonstrated that the proposed approach achieved low distance error, outperforming state-of-the-art research using LSTM algorithm [6]. This highlights the effectiveness of the PCA-WLR approach in the indoor position prediction task.

Table I presents the computational costs of the LSTM method and our proposed approach. Specifically, a single LSTM layer takes approximately 564.1396 seconds for training and model generation, with a testing time of 10.0848 seconds. In comparison, our approach only requires about 8.8814 seconds for train time and 0.0176 seconds for test time. Our approach resulted in approximately 98% reduction in training time and about 99% reduction in prediction time compared to the LSTM method [6]. Table II presents the proposed model parameters, which provides detailed information about the parameters of our proposed model. The table lists the number of parameters for each component of our hybrid model, including PCA with p=30 and WKNN with k=30, as well as the linear regression component. This enhancement greatly contributes to efficiency and computational resource savings. As a result, our proposed model achieves outstanding efficiency in both train time and prediction time compared to the state-of-the-art LSTM method.

TABLE I. Measurements of Computation Time

Model	Training Time(s)	Testing Time(s)
LSTM [6]	581.3599	10.1721
PCA-WLR	8.8814	0.0176

TABLE II. The Proposed Model Parameters

Model	The number of parameters
PCA	p=30
WKNN	k=30

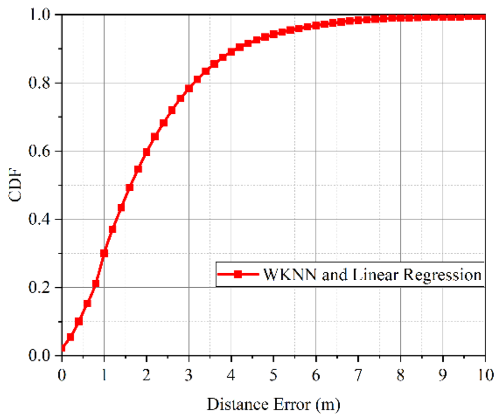

Fig. 3. The cumulative distribution functions.

The performance of the proposed algorithm is clearly depicted by the cumulative distribution function (CDF) illustrated in Fig. 3. The results demonstrate that the algorithm performs well, showcasing its effectiveness in predicting accurate distances. Furthermore, it's noteworthy that as the error increases beyond 4 meters, the rate of CDF increment slows down. This observation suggests that the algorithm's predictive capabilities are particularly strong for smaller errors, but it gradually becomes less sensitive to larger errors. In summary, the CDF analysis affirms the success of the proposed algorithm in accurately predicting distances. The gradual saturation of CDF for larger errors indicates the algorithm's ability to handle a range of errors with robustness and reliability.

This observation reveals the algorithm's capability to predict with lower probabilities for smaller errors, gradually becoming less sensitive to larger errors. This finding aligns perfectly with real-world scenarios, as the likelihood of accurate predictions tends to decrease with smaller distance errors.

The simulation results show that the dimensionality reduction from 448 features to 30 features using PCA also brought numerous benefits. PCA helped reduce computational and storage burden, eliminate unimportant features, enhance model interpretability and performance, and mitigate overfitting risks. As a result, this improvement led to increased prediction accuracy and better generalization of the model.

The combination of WKNN and Linear Regression also demonstrated high effectiveness in position prediction. WKNN leveraged information from the nearest data points to estimate the target position, while Linear Regression fine-tuned the predictions by learning linear relationships between features and positions. This synergy improved the accuracy of predictions and allowed the model to leverage spatial relationships between data points effectively.

With advantages such as computational efficiency, accuracy, ease of deployment, and high interpretability, the PCA-WLR approach proves to be a useful and effective solution for indoor position prediction. The choice between different methods depends on specific task requirements and data conditions, but in this case, the proposed method exhibited superior effectiveness and performance compared to the deep learning LSTM algorithms [6]. With a reduction in localization error in ranging 10.8% to 17.2%, these results mark a significant step forward in the research and application of position prediction solutions in indoor environments.

V. Conclusion

In this paper, we proposed a PCA-WLR approach for indoor position prediction. Our proposal demonstrated good performance, achieving a low localization error of 2.23 meters, which is higher to contemporary research that employed LSTM algorithms with errors ranging from 2.5 meters to 2.7 meters. Moreover, our proposal exhibited remarkable computational efficiency, significantly reducing training and prediction times. Compared to LSTM, our model achieved an approximately 98.47% reduction in training time, taking only 8.88 seconds, and a nearly 99.82% reduction in prediction time, requiring only 0.018 seconds. This accelerated computation not only enhances the efficiency of the model but also allows for real-time deployment, making it well-suited for applications requiring rapid indoor position predictions.

We believe that our research opens up promising avenues for further exploration and advancement in indoor position prediction methodologies. Future work may focus on optimizing the model further, exploring alternative dimensionality reduction techniques, or investigating the combination of other machine learning algorithms for even more accurate and efficient results. With our contribution, we anticipate that indoor position prediction will continue to progress and find practical applications in various domains, enhancing the overall user experience and enabling new possibilities for indoor navigation and location-based services.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Thi-Hang Duong: Methodology, Investigation, contributed to the simulation and analyzed the data. Manh Kha Hoang: Formal analysis, Supervision, proposed an idea. Anh Vu Trinh: Formal analysis, Supervision, proposed an idea. All authors had written the paper and approved the final version.

Funding

This work was supported by the Hanoi University of Industry (HaUI) under Grant No. 24-2023-RD/HĐ-ĐHCN.

References

[1] N. Singh, S. Choe, and R. Punmiya, "Machine learning based indoor localization using Wi-Fi RSSI fingerprints: An overview," *IEEE Access*, vol. 9, pp. 127150–127174, 2021.

[2] A. Nessa, B. Adhikari, F. Hussain, and X. N. J. I. A. Fernando, "A survey of machine learning for indoor positioning," vol. 8, pp. 214945–214965, 2020.

[3] X. Zhu *et al.*, "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," vol. 22, no. 4, pp. 2634–2657, 2020.

[4] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020.

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

[6] H.-Y. Hsieh, S. W. Prakosa, and J.-S. Leu, "Towards the implementation of recurrent neural network schemes for WiFi fingerprint-based indoor positioning," in *Proc. 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–5.

[7] B. Wang *et al.*, "A novel weighted KNN algorithm based on RSS similarity and position distance for Wi-Fi fingerprint positioning," vol. 8, pp. 30591–30602, 2020.

[8] B. Wang *et al.*, "A novel weighted KNN algorithm based on rss similarity and position distance for Wi-Fi fingerprint positioning," *IEEE Access*, vol. 8, pp. 30591–30602, 2020.

[9] B. Wang, Y. Zhao, T. Zhang, and X. Hei, "An improved integrated fingerprint location algorithm based on WKNN," in *Proc. 2017 29th Chinese Control and Decision Conference (CCDC)*, 2017, pp. 4580–4584.

[10] X. Song *et al.*, "A novel convolutional neural network based indoor localization framework with WiFi fingerprinting," *IEEE Access*, vol. 7, pp. 110698–110709, 2019.

[11] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. J. I. I. o. T. J. Reddy, "Recurrent neural networks for accurate RSSI indoor localization," vol. 6, no. 6, pp. 10639-10651, 2019.

[12] 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, 2020.

[13] J. Oh and J. Kim, "AdaptiveK-nearest neighbour algorithm for WiFi fingerprint positioning," *ICT Express*, vol. 4, no. 2, pp. 91–94, 2018.
[14] P. Dai, Y. Yang, M. Wang, R. J. W. C. Yan, and M. Computing, "Combination of DNN and improved KNN for indoor location fingerprinting," 2019.
[15] M. Nabati, H. Navidan, R. Shahbazian, S. A. Ghorashi, and D. J. I. S. L. Windridge, "Using synthetic data to enhance the accuracy of fingerprint-based localization: A deep learning approach," vol. 4, no. 4, pp. 1–4, 2020.
[16] X. Song *et al.*, "A novel convolutional neural network based indoor localization framework with WiFi fingerprinting," vol. 7, pp. 110698–110709, 2019.
[17] L. Zheng, B.-J. Hu, J. Qiu, and M. J. I. I. o. T. J. Cui, "A deep-learning-based self-calibration time-reversal fingerprinting localization approach on Wi-Fi platform," vol. 7, no. 8, pp. 7072–7083, 2020.
[18] G. M. Mendoza-Silva, P. Richter, J. Torres-Sospedra, E. S. Lohan, and J. J. D. Huerta, "Long-term WiFi fingerprinting dataset for research on robust indoor positioning," vol. 3, no. 1, p. 3, 2018.