A Deep Learning Approach for UWB-based Indoor UAV Localization Using TDoA and RSSI Data

Juwon Hong, Jiwon Suh, Mose Gu, and Jaehoon (Paul) Jeong
Department of Computer Science and Engineering, Sungkyunkwan University, Suwon, Republic of Korea
Email: {hongju2024, sjw6136, rna0415, pauljeong}@skku.edu

***Abstract*—Indoor localization for Unmanned Aerial Vehicles (UAV) is a fundamental prerequisite for autonomous operations in environments where GPS is unavailable. Although Ultra-Wideband (UWB) with Time Difference of Arrival (TDoA) provides a reasonable localization service to those UAVs, its accuracy is often limited by conventional filtering algorithms like the Kalman Filter (KF). The performance of these filtering algorithms is fundamentally limited by their dependency on simplified, predefined motion models, which are inadequate for tracking the agile and nonlinear movements of a UAV. To overcome this limitation, this paper introduces a deep learning framework centered on a Long Short-Term Memory (LSTM) network as a direct replacement for the Kalman Filter. The model is trained on a large-scale synthetic dataset, which includes sequences of TDoA and Received Signal Strength Indicator (RSSI) values from a UAV performing diverse flight patterns. By learning directly from the raw sensor stream, our model maps the complex temporal dependency to the UAV's positional coordinates without reliance on an explicit physical model. The empirical results from our simulations confirm that our data-driven approach yields a substantial improvement in localization accuracy and robustness over the traditional KF approach, presenting a viable and superior alternative for challenging indoor navigation tasks.**

***Index Terms*—TDoA, RSSI, Ultra-Wideband, Kalman filter, deep learning, LSTM, indoor localization**

I. Introduction

In recent years, the utilization of Unmanned Aerial Vehicles (UAV), commonly known as drones, has expanded rapidly in environments where human access is difficult or hazardous, such as disaster site exploration, large-scale warehouse automation, and precision inspection of indoor facilities [1]. The successful autonomous execution of these critical missions presupposes a precise and robust positioning capability. However, in environments such as indoors or urban canyons where Global Positioning System (GPS) signals are obstructed, satellite-based navigation becomes infeasible. This limitation makes Indoor Positioning Systems (IPS) as an essential alternative [2].

Among various indoor positioning technologies, Ultra-Wideband (UWB) has emerged as one of the most promising solutions, offering superior ranging accuracy based on its nanosecond-level temporal resolution and demonstrating strong resilience to multipath fading, a persistent challenge in indoor settings [3]. The Time Difference of Arrival (TDoA) technique, which utilizes the time difference of signals arriving at multiple fixed anchors from a mobile tag, is frequently employed in UWB systems as it obviates the need for strict time synchronization at the tag [4]. However, UWB signals are inevitably corrupted by noise stemming from factors like Non-Line-of-Sight (NLOS) conditions, where the signal path is obstructed by obstacles, and multipath interference.

To mitigate such noise and produce a stable trajectory estimate, model-based filtering techniques like the Kalman Filter (KF) and its variants, such as the Extended Kalman Filter (EKF), have been widely adopted [5]. Nevertheless, these approaches possess a fundamental limitation. The Kalman Filter assumes a linear motion model, such as Constant Velocity (CV) or Constant Acceleration (CA), which is inherently inconsistent with the agile and highly nonlinear dynamics of a UAV that frequently undergoes abrupt changes in acceleration and orientation. This discrepancy between the prediction model and the actual movement can severely degrade localization accuracy [6].

To overcome the limitations of these model-based filters, this paper proposes a data-driven, deep learning approach that learns motion patterns directly from raw sensor data. Specifically, we design an end-to-end positioning model using a Long Short-Term Memory (LSTM) network to completely replace the conventional Kalman Filter. The LSTM is a type of Recurrent Neural Network (RNN) renowned for its effectiveness in capturing complex temporal dependencies in time-series data [7]. The key contributions of this work are as follows:

- First, we propose an end-to-end LSTM-based localization framework that directly maps a sequence of raw UWB TDoA and RSSI data to the UAV's 3D position, without relying on a predefined physical model.
- Second, we developed a high-fidelity simulation environment to generate a large-scale dataset for training and evaluation, encompassing a wide variety of flight scenarios.
- Third, through extensive experiments on complex flight paths, we quantitatively demonstrate that the proposed model achieves an average of 18% improvement in positioning accuracy over the traditional Kalman Filter baseline.

In this paper, we initially assume a Line-of-Sight (LoS) environment to clearly validate the core performance of the proposed model. The remainder of this paper is organized as follows. Section II reviews related work. Section III details the proposed system architecture and LSTM model. Section IV describes the experimental setup along with dataset generation process. Section V presents performance evaluation through the comparative analysis of the experimental results. Finally, Section VI concludes this paper along with future work.

II. Related Work

UWB technology is widely recognized for its potential in high-precision indoor localization due to its excellent time-domain resolution and resilience to multipath fading. In UWB systems employing the TDoA technique, raw measurements are often corrupted by noise. To mitigate this and produce a smooth trajectory estimate for dynamic objects, filtering algorithms are traditionally employed, with the Kalman Filter (KF) and its variants being the most prominent choices. However, the performance of the Kalman Filter is fundamentally limited by its reliance on a predefined mathematical motion model, such as a CV or CA model. Such assumptions struggle to accurately capture the complex, nonlinear dynamics of an agile UAV. Furthermore, when UWB signals are temporarily lost, KF-based approaches that also utilize an Inertial Measurement Unit (IMU) suffer from rapid error accumulation, leading to a significant degradation in localization accuracy. This dependency on simplified models and vulnerability to signal outages necessitates a more robust, data-driven approach.

To address the limitations of model-based filters, researchers have increasingly turned to deep learning methods. Early applications focused on using neural networks to improve the quality of UWB measurements. For instance, [8] proposed DeepTAL, a LSTM based network designed to handle TDoA measurement errors and missing data in asynchronous localization systems. Their model learns the temporal patterns in TDoA sequences to predict and correct faulty or incomplete data points. Similarly, [9] developed a hybrid algorithm where an LSTM network predicts future TDoA values to correct real-time measurements affected by issues like clock drift. The corrected TDoA values are then fed into a separate Weighted K-Nearest Neighbors (WKNN) model to compute the final position. These studies demonstrate the effectiveness of LSTMs in enhancing the integrity of raw TDoA data, though they do not use the neural network for end-to-end position estimation itself.

More recent works have taken a step further by using RNNs to replace the Kalman Filter's role entirely. A highly relevant study by [10] proposed an RNN-based localizer that takes past UWB TDoA and IMU sensor data as input to predict future localization coordinates. Their key contribution is a mechanism to handle UWB signal outages: when the signal is lost, the model uses its own predicted location to generate augmented TDoA values, which are then fed back into the network. This allows the system to maintain stable tracking where a KF-based approach would fail, achieving a 31% lower localization error compared to the KF baseline. This trend is also visible in adjacent fields. For instance, [11] applied Convolutional Neural Networks (CNN) to estimate the ToA from raw acoustic emission signals in noisy environments,

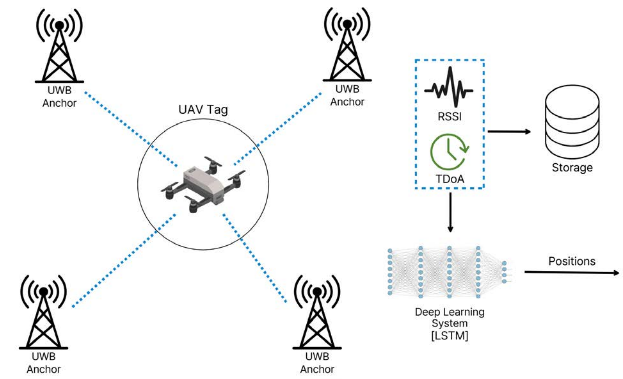

Fig. 1. Architecture of the Proposed LSTM-based UWB Localization System.

demonstrating a 10x gain in accuracy compared to traditional statistical methods. In the specific context of UAV navigation, deep learning has also been pivotal in enabling the fusion of multiple sensor modalities. [12] developed VIUNet, a framework that fuses Vision, IMU, and UWB data for indoor UAV localization, arguing that deep learning can learn and compensate for sensor biases more effectively than conventional methods.

Our work is positioned within this research landscape. While building on the demonstrated success of LSTMs for processing UWB time-series data, our approach distinguishes itself by focusing on a streamlined yet robust solution. Unlike multi-modal fusion systems such as VIUNet, we exclusively use UWB-derived data (TDoA and RSSI). In contrast to methods that use LSTMs only for data correction, we employ an end-to-end LSTM model to directly map the sequence of raw sensor readings to a 3D position, thereby replacing the Kalman Filter's function entirely. By directly comparing our model with a traditional KF baseline, we aim to demonstrate that a well-trained deep learning model can outperform conventional filters for complex 3D UAV trajectory tracking without the need for additional sensor types like IMU or cameras.

III. Proposed Method

A. System Architecture

The proposed system is an indoor positioning system designed to track the 3D trajectory of a UAV. The fundamental hardware configuration, as illustrated in Fig. 1, consists of two main components: multiple UWB anchors fixed at known positions within the environment and a single UWB tag mounted on the mobile UAV. This setup is specifically designed to utilize the TDoA measurement technique, which offers the practical advantage of not requiring precise time synchronization on the mobile tag [4].

The data processing pipeline begins when the UWB tag on the UAV periodically transmits signals. These signals are received by each of the stationary UWB anchors. For each transmission event, the system calculates the TDoA values, representing the difference in signal arrival times between a

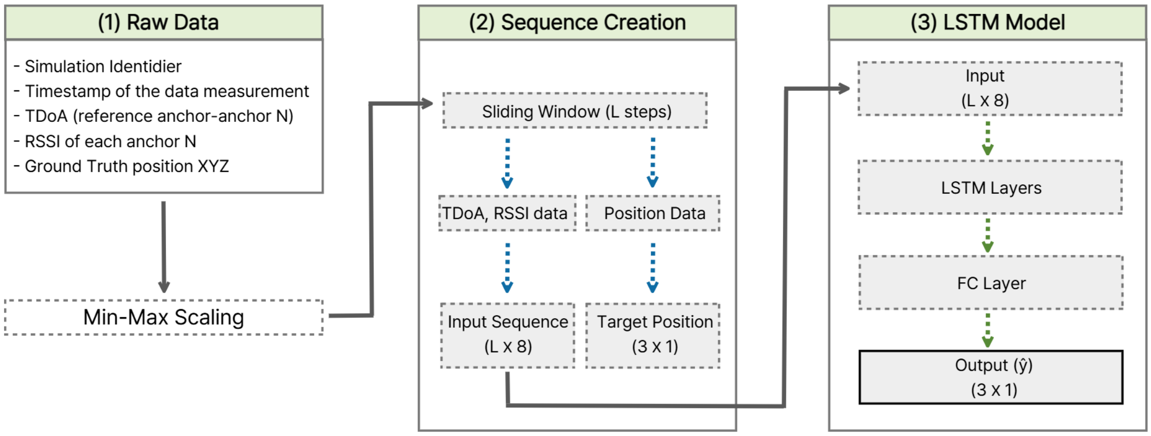

Fig. 2. Data Processing Pipeline and LSTM Model Architecture.

designated reference anchor and the other anchors. Simultaneously, the RSSI value is measured at each anchor.

This process yields a stream of raw sensor measurements containing both temporal (TDoA) and signal strength (RSSI) information. Instead of being filtered by a conventional model-based filter, this raw time-series data is directly fed into our proposed deep learning model. The model's objective is to learn the complex underlying patterns of the UAV's motion from this sequence of sensor data and output a robust estimate of the UAV's 3D position coordinates (x, y, z) at each time step.

B. Input Data Formulation

Our proposed model leverages the temporal context inherent in the UAV's movement by processing data in sequences rather than as individual, isolated data points. The core of our data-driven approach is the formulation of input data that preserves this time-series information.

First, at each discrete time step t, a feature vector $\mathbf{x}_t$ is constructed by concatenating the TDoA and RSSI values from all $N = 4$ anchors:

$$\begin{aligned}\mathbf{x}_t = [&TDoA_{1,t}, TDoA_{2,t}, TDoA_{3,t}, TDoA_{4,t},\\ &RSSI_{1,t}, RSSI_{2,t}, RSSI_{3,t}, RSSI_{4,t}]^T,\end{aligned} \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^8$.

To prepare the data for the LSTM model, we employ a sliding window technique as part of our data preprocessing pipeline. A fixed-length sequence of L consecutive feature vectors is extracted to form a single input sample, $\mathbf{X}_i$:

$$\mathbf{X}_i = [\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_{i+L-1}]. \tag{2}$$

Here, $\mathbf{X}_i \in \mathbb{R}^{L\times 8}$ is a matrix representing the sensor readings over a time window of length L, which is a key hyperparameter of our system. The corresponding ground truth label for this input sequence, $\mathbf{y}_i$, is the UAV's actual 3D position at the subsequent time step, $i + L$:

$$\mathbf{y}_i = [p_{x,i+L}, p_{y,i+L}, p_{z,i+L}]^T. \tag{3}$$

This structure tasks the model with predicting the future position based on the recent history of sensor measurements.

Prior to sequence creation, a crucial preprocessing step is applied. Both the input features (e.g., TDoA and RSSI) and the output targets (e.g., position coordinates) are independently normalized to a range of $[0, 1]$ using Min-Max scaling. This normalization is essential for the stable and efficient training of the neural network. The process ensures that no single feature dominates the learning process due to its scale. To maintain the integrity of our time-series data, sequence generation is performed strictly within the boundaries of each unique simulation run (identified by 'RunID'). This prevents the model from learning spurious patterns across disconnected flight trajectories.

C. LSTM-based Positioning Model

To overcome the limitations of predefined motion models inherent in traditional filtering methods, we propose a deep learning model based on a LSTM network. LSTM is a special type of RNN specifically designed to learn long-range temporal dependencies, making it highly effective for time-series data analysis [7]. Our model directly learns the complex spatiotemporal correlations from the sequence of UWB sensor data to estimate the UAV's trajectory.

The architecture of our proposed model, as depicted in Fig. 2, is implemented using the PyTorch framework and consists of the following layers:

- **Input Layer:** The model takes the sequence matrix $\mathbf{X} \in \mathbb{R}^{L\times 8}$ as input, where L is the sequence length (a hyperparameter set to 8 in our experiments) and 8 represents the number of input features (e.g., 4 TDoA values and 4 RSSI values).

- **Stacked LSTM Layers:** To capture the complex dynamics of the UAV's motion, we employ a stack of two LSTM layers with a hidden size of 128 neurons. Stacking LSTM layers allows the model to learn hierarchical temporal features, with the first layer learning basic temporal patterns and the second layer learning more abstract patterns from the output of the first. A dropout rate of 0.2 is applied between the LSTM layers to prevent overfitting by randomly deactivating a fraction of neurons during training, which enhances the model's generalization capability.
- **Fully Connected Layer:** The output of the second LSTM layer is a sequence of hidden states for each time step. We are interested in making a single position prediction after observing the entire sequence. Therefore, only the hidden state from the very last time step of the sequence (h_L) is passed to a fully connected (or dense) layer. This layer acts as a regressor, mapping the learned high-level feature representation from the LSTM's final hidden state to the desired 3D output space.
- **Output Layer:** The final layer outputs a vector $\hat{\mathbf{y}} \in \mathbb{R}^3$, which represents the model's estimate of the UAV's 3D position coordinates $(\hat{p}_x, \hat{p}_y, \hat{p}_z)$.

For the model's training, we use the Mean Squared Error (MSE) as the loss function, which measures the average squared difference between the estimated positions and the ground truth positions. The model's weights are optimized using the Adam optimizer, an efficient stochastic gradient descent algorithm, with a learning rate of 0.0008.

IV. Experimental Setup

To rigorously evaluate the performance of our proposed LSTM-based model against the traditional Kalman Filter approach, we designed a comprehensive simulation-based experiment. This section details the methodology for generating a large-scale dataset, the baseline model used for comparison, and the metrics for performance evaluation.

A. Dataset Generation

A key requirement for training a robust deep learning model is a large and diverse dataset that accurately reflects the complexities of the target environment. To this end, we developed a sophisticated simulation environment using MATLAB and its UAV Toolbox to generate synthetic UWB sensor data from a wide variety of UAV flight paths [13].

The simulation environment was configured with four stationary UWB anchors placed at known coordinates within a 3D space. The core of our data generation process involved executing 1,000 unique simulation runs. For each run, a new, complex flight trajectory for the UAV was procedurally generated. This was achieved by first defining a set of random waypoints within the simulation boundaries. Subsequently, a greedy algorithm was employed to reorder these waypoints, creating a more realistic and dynamically complex flight path than a simple traversal of random points would allow. The UAV was programmed to navigate this path at an average speed, ensuring varied velocity and acceleration profiles.

During each simulation, as the UAV traversed its trajectory, we logged data at every valid UWB signal transmission event. Our custom UWB sensor model simulates transceivers compliant with the IEEE 802.15.4z standard, operating at a center frequency of 6.5 GHz in High Pulse Repetition Frequency (HPRF) mode [14]. This model was used to simulate signal propagation, calculating the TDoA and RSSI values at each anchor with the inclusion of realistic noise factors such as random fading effects. For each valid data point, we recorded the following information:

- **RunID:** A unique identifier for each of the 1,000 simulation runs.
- **Timestamp:** The simulation time at which the measurement was taken.
- **Ground Truth Position (x, y, z):** The true 3D coordinates of the UAV at the moment of transmission, obtained directly from the simulation engine.
- **TDoA Values:** A set of four TDoA measurements, one for each anchor relative to the reference anchor.
- **RSSI Values:** A set of four RSSI measurements, indicating the signal strength received at each anchor.

This process resulted in a comprehensive dataset containing thousands of data points, encapsulating a wide range of flight dynamics. The final dataset was exported as a single CSV file, which served as the basis for training and evaluating both the proposed LSTM model and the baseline Kalman Filter.

B. Baseline Model: Kalman Filter

To validate the effectiveness of our proposed deep learning approach, we selected the KF as the baseline model for performance comparison. The Kalman Filter is a traditional and widely-adopted algorithm for state estimation and trajectory smoothing in noisy dynamic systems, making it a standard benchmark in UWB-based localization tasks [5].

Our implementation of the Kalman Filter is designed to process the UWB TDoA measurements and produce a smoothed 3D position estimate. The specific configuration of our baseline model is as follows:

- **State Vector:** The state of the UAV at each time step k is represented by a 6-dimensional vector $\mathbf{x}_k$, which includes its 3D position and 3D velocity components:

$$\mathbf{x}_k = [p_x, p_y, p_z, v_x, v_y, v_z]^T_. \tag{4}$$

- **Motion Model:** For the filter's prediction step, we assume a linear, CV motion model. This model predicts the UAV's next state by assuming it will continue to move at its current velocity. This reliance on a predefined physical model is a key characteristic of the KF approach and a primary point of contrast with our data-driven LSTM model.
- **Measurement Model:** The direct measurements from the UWB system are the TDoA values, which have a non-linear relationship with the UAV's position. Therefore, at

each time step, we first compute a raw position estimate from the TDoA values using a nonlinear least squares optimization algorithm. This raw 3D position serves as the measurement input $\mathbf{z}_k$ for the Kalman Filter's update step.

This Kalman Filter setup represents a robust and conventional method for UWB localization. It serves as a strong baseline to demonstrate the advantages of our proposed LSTM model, particularly in its ability to capture the complex, nonlinear dynamics of UAV flight without being constrained by a predefined motion model.

V. Performance Evaluation

To conduct a fair and direct comparison between the proposed LSTM model and the baseline Kalman Filter, we performed a series of 50 simulations, each with a unique, randomly generated flight path. For every one of these 50 distinct trajectories, both the baseline model and our proposed deep learning model were tasked with estimating the UAV's position. This one-to-one evaluation on identical paths ensures that the performance differences can be directly attributed to the models' capabilities rather than variations in the flight dynamics. During each simulation, the positioning error for both methods was meticulously recorded for quantitative analysis.

A. Evaluation Metrics

The positioning accuracy of each model is quantitatively assessed using two standard statistical error metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). These metrics measure the difference between the model's estimated position $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ and the ground truth position (x_i, y_i, z_i) for a total of N data points.

The primary metric is the **Root Mean Square Error (RMSE)**, which provides a measure of the error magnitude and is particularly sensitive to large, infrequent errors due to the squaring term. It is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2)}. \tag{5}$$

Additionally, the **Mean Absolute Error (MAE)** is employed, which calculates the average of the absolute error magnitudes (Euclidean distance) across all data points. This metric provides a more direct interpretation of the average error size and is defined as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2}. \tag{6}$$

These two metrics collectively offer a comprehensive view of each model's accuracy, allowing for a robust comparison of their overall performance.

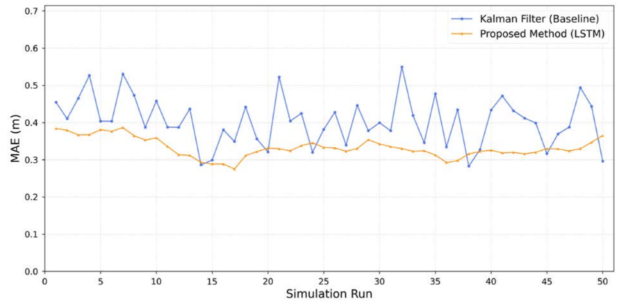

Fig. 3. Comparison of Mean Absolute Error (MAE) for Each Simulation Run.

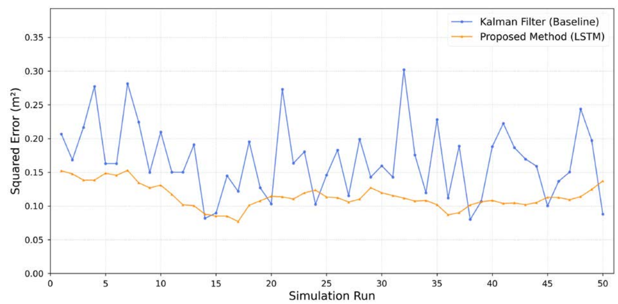

Fig. 4. Comparison of Squared Error for Each Simulation Run.

B. Testing Result

The performance of the proposed LSTM model and the baseline Kalman Filter was evaluated across all 50 unique simulation runs. Figs. 3 and 4 provide a detailed, run-by-run comparison of the MAE and the Squared Error, respectively. Both figures reveal a consistent visual pattern: the proposed LSTM model (orange line) demonstrates significantly lower positioning errors compared to the baseline Kalman Filter (blue line) across most randomly generated trajectories. This consistent performance advantage highlights the enhanced accuracy and robustness of our approach across diverse flight scenarios.

For a quantitative summary of the overall performance, the aggregated statistics are presented in Table I. The numerical data reinforces the visual evidence from the graphs. Our LSTM model achieved an Average MAE of 0.3324 m, representing a 17.8% improvement over the Kalman Filter's 0.4042 m. A similar, more pronounced trend is observed in the RMSE, where the LSTM model (0.3391 m) outperformed the baseline (0.4183 m) by 18.9%.

TABLE I
Quantitative Comparison of Positioning Error

Metric	Kalman Filter (Baseline)	LSTM Model
Average MAE [m]	0.4042	**0.3324**
Minimum MAE [m]	0.2828	**0.2752**
Maximum MAE [m]	0.5497	**0.3865**
RMSE [m]	0.4183	**0.3391**

Furthermore, the analysis of best and worst-case scenarios highlights the robustness of our data-driven approach. In the worst-performing run, the LSTM model's Maximum MAE was only 0.3865 m, which is considerably lower than the Kalman Filter's worst-case error of 0.5497 m. The LSTM model also achieved a better best-case performance, with a Minimum MAE of 0.2752 m compared to the Kalman Filter's 0.2828 m. Both the detailed visual evidence and the aggregate statistical results strongly support the conclusion that the proposed LSTM model provides not only higher average accuracy but also more reliable and robust localization service for complex UAV trajectories than the conventional Kalman Filter approach.

VI. Conclusion

In this paper, we proposed and validated a Long Short-Term Memory (LSTM)-based deep learning model for the 3D indoor localization of Unmanned Aerial Vehicles (UAV) using raw Ultra-Wideband (UWB) Time Difference of Arrival (TDoA) and Received Signal Strength Indicator (RSSI) data. Through comprehensive experiments conducted on a large-scale, procedurally generated simulation dataset, we demonstrated that our proposed model achieves significantly higher accuracy and robustness than the traditional Kalman Filter approach. The results were particularly compelling for complex flight trajectories, where the data-driven model excelled at tracking nonlinear dynamics. This paper suggests that deep learning-based, data-driven methods are effective for overcoming the inherent limitations of conventional filters that rely on predefined physical motion models.

Building on these promising results, future research will proceed in several key directions. The immediate next step is to validate the proposed model in a physical environment using real UWB hardware and UAVs, which will involve addressing more challenging real-world conditions, including Non-Line-of-Sight (NLOS) signal propagation. Furthermore, we will explore and compare the performance of our current architecture with the performance of other advanced sequence-to-sequence models, such as Gated Recurrent Units (GRUs) and the Transformer, to potentially achieve further improvement in both localization accuracy and computational efficiency.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No. 2023R1A2C2002990), and also by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2024-00398199). Note that Jaehoon (Paul) Jeong is the corresponding author.

References

[1] M. Erdelj and E. Natalizio, "Uav-assisted disaster management: Applications and open issues," in *2016 International Conference on Computing, Networking and Communications (ICNC)*, 2016, pp. 1–5.

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[3] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrami, M. A. Al-Ammar, and H. S. Al-Khalifa, "Ultra wideband indoor positioning technologies: Analysis and recent advances," *Sensors*, vol. 16, no. 5, 2016.

[4] J. Tiemann, F. Eckermann, and C. Wietfeld, "Atlas - an open-source tdoa-based ultra-wideband localization system," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2016, pp. 1–6.

[5] P. Jiang, C. Hu, T. Wang, K. Lv, T. Guo, J. Jiang, and W. Hu, "Research on a visual/ultra-wideband tightly coupled fusion localization algorithm," *Sensors*, vol. 24, no. 5, 2024.

[6] Y. Li, M. Jiu, Q. Sun, and Y. Wang, "An improved target tracking algorithm based on extended kalman filter for uav," in *2018 IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP)*, 2018, pp. 435–437.

[7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[8] Y. Xue, W. Su, H. Wang, D. Yang, and Y. Jiang, "Deeptal: Deep learning for tdoa-based asynchronous localization security with measurement error and missing data," *IEEE Access*, vol. 7, pp. 122 492–122 502, 2019.

[9] H. Wang, X. Wang, Y. Xue, and Y. Jiang, "Uwb-based indoor localization using a hybrid wknn-lstm algorithm," in *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, vol. 1, 2020, pp. 1720–1725.

[10] S. Bhattacharya and J. Choi, "Rnn-based robust smartphone indoor localization on ultra-wideband dl-tdoa," in *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2023, pp. 500–505.

[11] F. Zonzini, D. Bogomolov, T. Dhamija, N. Testoni, L. De Marchi, and A. Marzani, "Deep learning approaches for robust time of arrival estimation in acoustic emission monitoring," *Sensors*, vol. 22, no. 3, 2022.

[12] P.-Y. Kao, H.-J. Chang, K.-W. Tseng, T. Chen, H.-L. Luo, and Y.-P. Hung, "Viunet: Deep visual–inertial–uwb fusion for indoor uav localization," *IEEE Access*, vol. 11, pp. 61 525–61 534, 2023.

[13] T. M. Inc., "Matlab version: (r2025a)," Natick, Massachusetts, United States, 2025. [Online]. Available: https://www.mathworks.com

[14] The MathWorks Inc., "Uwb ranging using IEEE 802.15.4z," 2025. [Online]. Available: https://kr.mathworks.com/help/comm/ug/uwb-ranging-using-ieee-802.15.4z.html