Deep Bidirectional LSTM-Based Sensor Fusion for 3D Localization in Dynamic Environments

Tehmina Bibi*, Anselm Köhler†, Jan-Thomas Fischer†, and Falko Dressler*
*School of Electrical Engineering and Computer Science, TU Berlin, Germany
†Austrian Research Centre for Forests, Austria
{bibi,dressler}@tkn.tu-berlin.de, {anselm.koehler,jt.fischer}@bfw.gv.at

***Abstract*—Accurate motion tracking in snow avalanche environments remains challenging. Recently, ultra-wideband (UWB)-based ranging systems have been shown to offer centimeter-level accuracy under ideal conditions. However, their performance degrades significantly in dynamic, obstructed snow to multipath and signal attenuation. This can be compensated using multisensor fusion based on inertial measurement unit (IMU) data. To address these limitations, we propose a bidirectional LSTM (Bi-LSTM) network for robust 3D localization. The proposed model leverages time-synchronized sliding windows to learn spatiotemporal correlations between high-frequency IMU dynamics and low-frequency UWB updates. Unlike traditional Kalman filters, which degrade under nonlinear motion and intermittent UWB outages, the Bi-LSTM model adaptively learns to weight each sensor modality while preserving long-term motion context. We evaluate our approach using real-world field data, using a RTK-GNSS as ground truth for validation. The proposed method achieves a 3D localization root mean square error of 0.31 m, representing a 68 % improvement over adaptive Kalman filter fusion and a 72 % gain over UWB-only localization. These results highlight the efficacy of deep recurrent sensor fusion for localization in dynamic, extreme environments.**

***Index Terms*—Localization, tracking, sensor fusion, bidirectional LSTM, ultra-wideband (UWB), inertial measurement unit (IMU), Kalman filter, deep learning**

I. Introduction

Localization in outdoor environments cannot be reliably achieved using a single positioning sensor, particularly under dynamic conditions. Although global navigation satellite system (GNSS) enables globally consistent real-time positioning, its accuracy and reliability often deteriorate in dynamic or obstructed environments [1]. This limitation becomes especially critical in extreme scenarios like snow avalanches [2], where rugged terrain, unpredictable motion, signal attenuation, and deep snow layers pose additional challenges. The challenge becomes even more severe in GNSS-denied environments, where satellite signals are blocked, rendering real-time localization infeasible.

The AvaRange project addresses these challenges by developing a distributed, particle-based tracking system aimed to investigate internal avalanche flow dynamics [3], [4]. This system employs ultra-wideband (UWB) ranging to localize mobile tags embedded within the snowpack in real-time. The idea is to deploy fixed anchors across large alpine regions, while mobile nodes perform time of flight (ToF) ranging combined with inertial measurement unit (IMU) samples to estimate positions. Field experiments conducted near Innsbruck Nordkettenbahn featuring controlled tests using a cable car to validate the feasibility of UWB tracking under avalanche-like conditions. Results as reported in this paper show that UWB-based ToF-based ranging improves accuracy over standard radio frequency (RF) techniques such as GNSS. However, robust 3D trilateration requires simultaneous visibility to at least four anchors, which is often unachievable in complex terrain. Additionally, localization accuracy suffers with increased anchor spacing or suboptimal anchor geometry due to geometric dilution of precision (GDOP).

Findings from the conducted experiments also reveal that, although GDOP effectively indicates geometric sensitivity to noise, it does not accurately capture the true quality of measurements or filtering performance often resulting in overly conservative error estimates.

In dynamic environments, localization relies primarily on two methods: geometric methods and fingerprinting methods. Each offers unique advantages and trade-offs based on environmental complexity, required accuracy, and computational resources. Geometric localization techniques, such as trilateration and triangulation, estimate a target's position using measurable physical quantities, such as distances or angles, between the mobile tag and known anchor points. These methods are computationally efficient, sensor-agnostic, and well-suited for real-time applications. However, their performance tends to degrade in complex environments due to multipath interference or obstructed line-of-sight (LOS) conditions [5]. In contrast, fingerprinting techniques rely on a pre-collected database of signal characteristics such as received signal strength indicator (RSSI) and channel impulse response (CIR) measured at known locations during an offline phase. During operation, real-time measurements are compared against this database to infer the target's position. Even though fingerprinting can offer high accuracy in cluttered or non-line-of-sight (NLOS) environments, it requires labor-intensive calibration, and frequent retraining, and is often sensitive to environmental changes [6].

In summary, geometric methods based on ToF techniques offer strong potential for UWB-based positioning systems; however, several challenges remain unresolved. These include poor anchor visibility, environmental variability, signal reflections, sensor errors, and the need for costly setup and tuning. Such problems make traditional positioning systems less reliable in complex outdoor environments. Many existing correction methods depend on dense anchor deployments or computationally intensive post-processing techniques [7]–[10].

As an alternative, researchers have explored inertial navigation systems (INS), particularly those utilizing IMUs. A typical IMU includes a three-axis accelerometer and gyrometer, providing high-rate measurements of acceleration, angular velocity, and orientation [11]–[13]. Strapdown inertial navigation systems (SINS) integrate this data to estimate position without requiring external references. However, accuracy degrades over time due to sensor errors both deterministic (e.g., bias, scale-factor, axis misalignment, temperature sensitivity) and stochastic (e.g., drift, random noise) [14], [15]. While deterministic errors can often be mitigated through calibration, stochastic errors are inherently unpredictable, causing cumulative drift.

The complementary characteristics of UWB and IMU sensors make sensor fusion an effective solution for improving localization performance. Traditionally, Bayesian filtering techniques such as the popular Kalman filter (KF) [16], extended Kalman filter (EKF) [17], unscented Kalman filter (UKF) [18], and particle filter (PF) [19] have been employed for fusing heterogeneous sensors. However, these approaches suffer from limitations including linearization errors, high computational cost, and suboptimal performance in nonlinear and non-Gaussian settings [20].

In recent years, deep learning-based fusion methods have emerged as powerful alternatives, capable of learning complex nonlinear relationships directly from data. For example, convolutional neural network (CNN)-based architectures have been used for classifying UWB LOS and NLOS conditions [21], and hybrid models integrating double extended Kalman filter (DEKF) with neural networks have demonstrated improved robustness [22]. Earlier data-driven methods, such as multilayer perceptron (MLP) [23] and support vector machines (SVMs) [24], offered limited compensation for sensor errors. Despite this, these methods could not model long-term temporal dependencies, which are critical for reliable navigation during GNSS outages. This gap has been addressed by recurrent architectures like long short-term memory (LSTM) networks [25], gated recurrent units (GRUs) [26], [27], and attention-based models [28], offering strong adaptability in dynamic environments.

To improve the 3D localization accuracy in avalanche environments based on UWB and IMU measurements, this paper explores both traditional Bayesian filtering and deep learning-based fusion methods. In particular, a bidirectional LSTM (Bi-LSTM) model is developed to fuse time-synchronized IMU accelerometer readings and trilaterated UWB position estimates. It learns to predict 3D motion increments, which are integrated over time during inference to reconstruct the full trajectory, ensuring temporal continuity and coherent motion estimation. The model processes input data using a sliding window to capture temporal motion patterns. During training, a custom weighted mean squared-error (WMSE) loss is applied to emphasize vertical (z-axis) accuracy, where UWB readings are typically less reliable, while also encouraging smooth and physically consistent trajectory predictions. Our proposed Bi-LSTM model achieves localization accuracy with a 72.2% improvement over UWB-only trilateration and 68.0% over an adaptive Kalman filter (AKF) baseline. These results demonstrate that the proposed deep learning-based fusion approach significantly outperforms conventional methods, offering a robust solution for motion tracking in highly dynamic and unstructured environments such as snow avalanches.

Our main contributions can be summarized as follows:

- We propose a Bi-LSTM model to fuse time-synchronized IMU accelerometer readings and trilaterated UWB position estimates;
- we trained and validated the model based on experimental data from a measurement campaign in the Austrian Alps;
- and we evaluated the proposed system in comparison with multiple baselines using a high-precision RTK GNSS system to provide ground truth.

II. Related Work

A. Radio-based UWB Localization

UWB-based ranging and localization has proven successful in delivering centimeter-level positioning accuracy. However, this accuracy is highly dependent on LOS conditions between the mobile node and the anchors. NLOS signal propagation and poor anchor geometry commonly referred to as GDOP can significantly degrade ToF measurements and overall localization performance [29].

Poulose et al. [30] introduced an UWB-based position estimator using an EKF for improved indoor localization in a multi-anchor setup. Their method processes time of arrival (ToA) data from multiple anchors to address system nonlinearities and was evaluated using MATLAB simulations in a 2D environment under both LOS and NLOS conditions. The results demonstrated that increasing the number of anchors enhanced accuracy while reducing computational complexity. Compared to traditional techniques such as linearized least squares estimation (LLSE), weighted centroid estimation (WCE), and maximum likelihood estimation (MLE), the EKF-based approach achieved accuracy within a ±50 cm margin. However, its use in dynamic scenarios is constrained by its dependence on predefined noise models, limiting its adaptability in real-world conditions.

Van Herbruggen et al. [31] proposed two heuristic anchor selection algorithms for two-way ranging (TWR)-based UWB indoor positioning systems to improve accuracy and update rates. The first method selects anchors based on link quality and GDOP, while the second is a lightweight, real-time algorithm suitable for deployment on low-power UWB tags. Both methods have been validated in industrial settings using a Kalman filter and achieved localization accuracies in the order of 15–20 cm. Although promising, the non-constrained algorithm requires continuous communication with all anchors and is computationally intensive, restricting its scalability. Moreover, both algorithms rely on static link quality metrics and lack mechanisms to account for motion dynamics or adapt to temporal changes which is an essential aspect for reliable localization in dynamic environments.

B. IMU-based and Multisensor Fusion

Multisensor fusion approaches combining IMUs with other modalities such as vision [32], [33], LiDAR [34], or UWB [35], [36] have gained popularity for improving localization accuracy. Among these, UWB/IMU fusion stands out as a cost-effective and compact solution capable of maintaining sub-meter accuracy even under NLOS conditions. However, it is still susceptible to multipath effects, IMU drift, and sensor noise [37], [38]. Classical filters such as KF, EKF, and UKF have been extensively applied [39], [40]; nonetheless, their reliance on linear models limits usefulness in nonlinear and non-Gaussian settings.

Hybrid approaches that integrate vision or LiDAR with UWB/IMU [41], [42] further improve robustness but incur high computational costs and are sensitive to environmental factors such as lighting and occlusions. Neurauter et al. [43] proposed two fusion algorithms using low-cost IMU gyrometer and magnetometer readings later merged with state-of-the-art Madgwick's algorithm for correcting orientation estimation, and tested on real snow avalanche datasets. The reliance on magnetometer data adds sensitivity to magnetic disturbances, thereby limiting its resistance to magnetically sensitive environments.

C. Learning-based Sensor Fusion Techniques

Data-driven learning models have been proposed to improve sensor fusion, primarily by modeling error distributions and confidence estimation. For instance, Tommingas et al. [44] introduced an extreme gradient boosting (XGBoost)-enhanced fusion method combining UWB and GNSS to manage indoor-outdoor transitions by predicting sensor confidence levels. However, their approach depends on GNSS availability and hand-crafted features, which limit scalability.

Deep neural networks have demonstrated strong potential in enhancing localization. For example, CNN-LSTM architectures [45] have been used for NLOS detection and correction, even though they still suffer from issues like sensor drift and synchronization challenges. The tightly coupled factor graph technique [46] models probabilistic sensor relationships and achieves high accuracy in controlled conditions, yet it is computationally expensive and sensitive to modeling errors.

Another neural network model, Bi-LSTM has been successfully applied in related fields such as human activity recognition and tracking [47]. Despite its capability to model bidirectional temporal dependencies and manage complex motion dynamics, it remains unexplored in the context of UWB/IMU fusion.

D. Proposed Bi-LSTM-based UWB/IMU Fusion

In contrast, this work presents a loosely coupled Bi-LSTM-based UWB/IMU fusion framework designed for GNSS-denied and dynamic environments. Instead of relying on GNSS inputs, hand-crafted filters, or explicit signal classification, our model learns temporal motion patterns directly from sequential sensor data. Compared to traditional Bayesian filtering methods, the neural network approach demonstrates higher accuracy and adaptability. It generalizes well across diverse and complex real-world scenarios while maintaining robustness against sensor drift, noise, GDOP, and NLOS conditions. This makes it especially suitable for accurate 3D trajectory estimation in extreme environments.

III. System Model

The system design combines UWB-based ranging with inertial sensing, thereby creating a multi-sensor dataset suitable for both traditional filtering and deep learning-based fusion techniques. As illustrated in Figure 1, the system follows a modular pipeline that begins with synchronized data collection from the nodes and anchors, followed by pre-processing of the data sets, sensor fusion, and final trajectory evaluation. The system workflow is divided into four main stages: (1) experimental setup and field data acquisition, (2) sensor-specific preprocessing comprising IMU calibration, motion reconstruction, frame transformation, and UWB-based localization, (3) multi-sensor fusion using AKF and Bi-LSTM neural networks, and (4) evaluation using GNSS-based ground truth. Each stage of the system is described in detail in the following.

A. Experimental Setup and Data Collection

For this work, we have used data collected in a field experiment in the Austrian Alps [48]. The experimental setup included an AvaNode mounted to a cable car counterweight. The cable car of Innsbruck Nordkettenbahn moved along a potential avalanche path emulating the actual avalanche scenario. This experiment was designed to emulate avalanche-relevant motion by tracking a cable car along a controlled, repeatable trajectory that approximates real avalanche paths. A sensor payload mounted on the cable car counterweight captured synchronized measurements from multiple sensor modalities: an UWB module, an IMU, and a high-precision RTK GNSS system.

1) AvaNode and AvaAnchors: The experimental system comprises two core device types: the AvaNode [49], which serves as the mobile sensing platform, and multiple AvaAnchors, statically deployed throughout the test area to facilitate position estimation [48]. The AvaNode, representing the tracked object, simulates the motion of a snow particle within an avalanche. The important sensing components include:

- DW1000 UWB transceiver for ToF-based ranging, interfaced with an Adafruit Feather M0 microcontroller. The DW1000 supports timestamping precision down to 15 ps, enabling highly accurate distance estimation.
- MPU9250 IMU, consisting of a 3-axis accelerometer and gyroscope, enclosed within a 3D-printed cube (edge length 100 mm) that aligns the sensor coordinate frame with the cube's geometric center.

An Emlid RTK GNSS unit is used for high-accuracy position ground truth. IMU data was logged at 400 Hz to an onboard SD card via the Feather M0 microcontroller. In contrast, UWB ranging was performed at a frequency of 20 Hz. Although the AvaNode system includes additional components such as GNSS receivers and a recovery mechanism, only data from the UWB

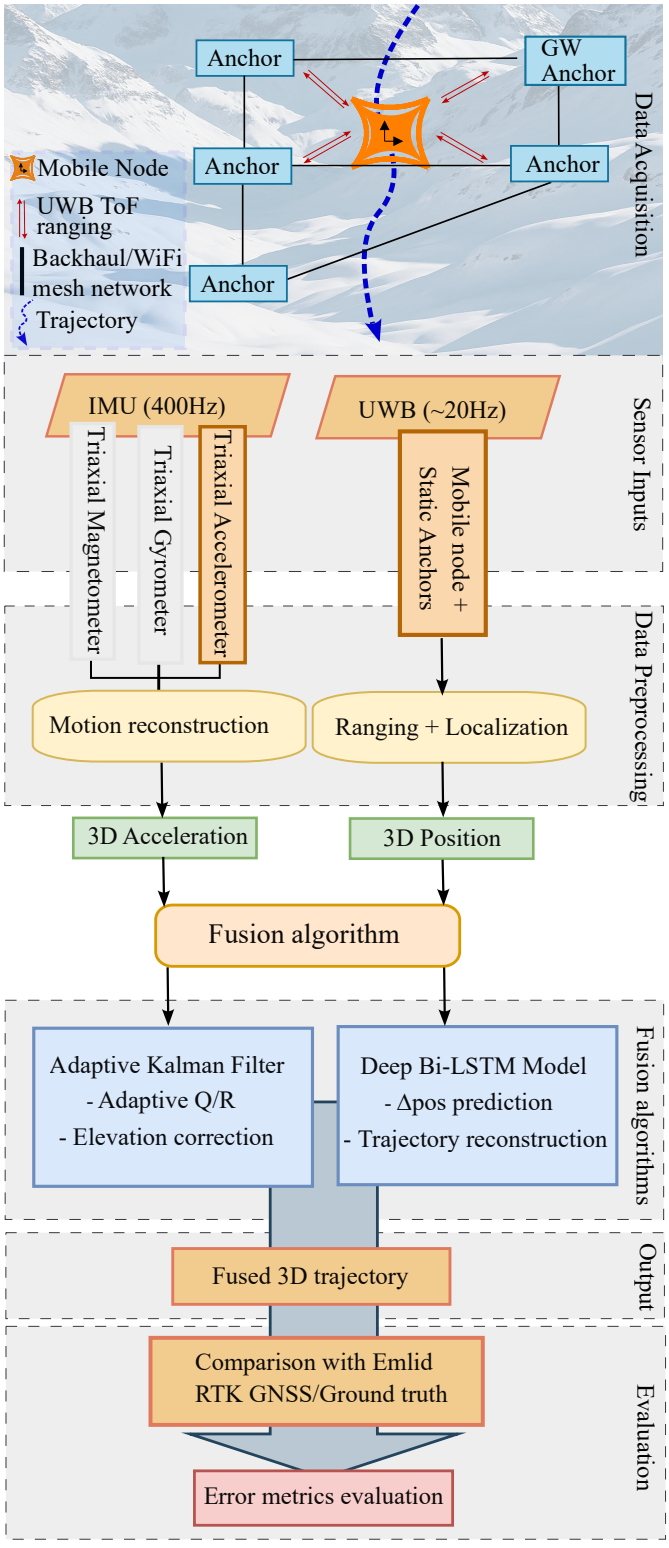
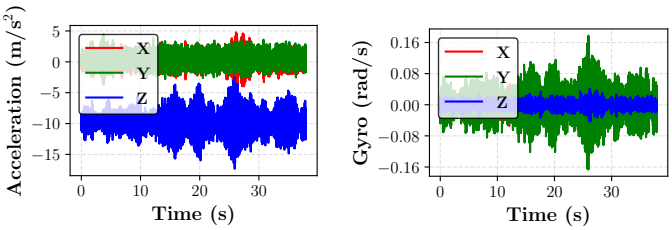


Figure 1. System architecture for UWB/IMU system using a deep Bi-LSTM model for 3D positioning and tracking in avalanches



(a) Accelerometer (b) Gyrometer

Figure 2. Calibrated accelerometer and gyrometer measurements along X-, Y-, and Z-axes.

and IMU subsystems are utilized in this work for motion reconstruction and localization analysis. The AvaAnchors, deployed at six fixed locations along the cable car track, also use DW1000 UWB transceivers. Their positions were precisely georeferenced using the real-time kinematic (RTK) GNSS system, ensuring accurate trilateration. These anchors formed a wireless mesh WiFi-based network, with at least one anchor configured as a gateway to manage the network, coordinate ranging schedules, and relay data to a backend server. This setup provided a robust framework for capturing high-resolution motion data under controlled conditions, enabling the evaluation of sensor fusion methods for 3D trajectory reconstruction in avalanche-like environments.

B. IMU Calibration and Motion Reconstruction

IMU calibration and motion reconstruction described in this work is based on the methodology proposed by Neurauter and Gerstmayr [50] and Winkler et al. [51], which models sensor errors, computes orientation from angular velocity, and enables gravity-compensated integration of accelerometer data. The calibration of IMU sensors is essential to reduce deterministic measurement errors such that accelerometer reflects only inertial and gravitational components, while the gyrometer is correctly tracking angular motion without introducing systematic drift.

The calibrated inertial measurements are depicted in Figure 2. The IMU sensor has been mostly subject to translational motion with less rotational activity. The accelerometer data shows consistent signals along the X- and Y-axes, while the Z-axis is dominated by the gravitational component, centered around $-9.81\ \mathrm{m/s^2}$, as expected. The gyrometer measurements have been close to zero across all axes, with minimal fluctuations which indicate absence of significant rotational motion.

The calibrated IMU data is then evaluated to recover the 3D trajectory of the mobile system using inertial sensor data, specifically translational acceleration and angular velocity. Motion reconstruction using IMU data refers to the process of estimating the full kinematic state (position, velocity, orientation) of the mobile node by integrating accelerometer and gyroscope measurements over time. Orientation estimation from calibrated gyroscope measurements involves computation of rotation matrices which help mapping vectors from local frame (body) to global frame of reference. Then, linear motion is computed by the transformation of calibrated accelerometer

readings to global coordinate frame using rotation matrix, followed by the integration of calibrated accelerometer signals. The gravity vector is subtracted to obtain translational acceleration, and the resulting global acceleration is numerically integrated using the explicit Euler method to obtain velocity and position. The computed gravity-compensated translational acceleration and motion data form a globally consistent 3D trajectory, suitable for comparison with external references such as UWB or GNSS systems.

C. UWB Ranging and Localization

For UWB-based distance estimation, the alternative double-sided two-way ranging (AltDS-TWR) method is adopted. The AltDS-TWR method enhances ranging accuracy by exchanging three messages between the mobile node and anchor, thereby reducing timing uncertainty and alleviating clock drift without requiring tight synchronization. This approach is robust to clock drift and does not require symmetric delays, making it well-suited for asynchronous systems [52]. The time-of-flight (T_f) is calculated as:

$$T_f = \frac{R_a R_b - D_a D_b}{2(R_a + D_a)}, \tag{1}$$

where R_a and R_b denote round-trip delays measured at each device, and D_a, D_b are the respective known reply delays. Distances d_i can be calculated from time-of-flight (T_f) measurements as:

$$d_i = c \cdot T_f, \tag{2}$$

where c denotes the speed of light.

The position of the mobile node is estimated through multilateration formulated by non-linear optimization. The objective is minimization of mean squared error (MSE) between the measured distances and the Euclidean distances to a set of fixed anchor nodes. Let $\mathbf{a}_i = (x_i, y_i, z_i)$ denote the coordinates of anchor i, and let d_i represent the corresponding measured distance. The estimated position $\mathbf{p} = (x, y, z)$ is obtained by solving the following optimization problem:

$$\text{MSE}(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^{N} \left(\|\mathbf{p} - \mathbf{a}_i\| - d_i \right)^2 \tag{3}$$

where N is the number of anchors providing distance measurements. A minimum of four anchors is required to obtain a unique solution in three-dimensional space.

Figure 3 compares the UWB-only trajectory with the RTK GNSS baseline. An increasing deviation of up to 4 m is observed over time, particularly along the vertical (Down) axis, where GDOP effects are more significant.

D. Sensor Fusion Using Adaptive Kalman Filter

We used an AKF as our baseline method for fusing data from UWB and IMU sensors. AKF extends functionality of a standard KF by dynamic adjustment of measurement noise model to enhance robustness under varying sensor noise and uncertainties. AKF uses standard prediction and correction (update) steps to estimate 3D position, velocity, and acceleration. The state vector can be defined as:

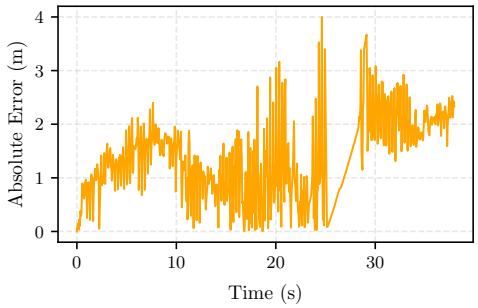


Figure 3. Relative error computed between UWB-only trilaterated positions and RTK GNSS ground truth along Z-axis.

$$\mathbf{x}_k = \begin{bmatrix} x_k & y_k & z_k & \dot{x}_k & \dot{y}_k & \dot{z}_k & \ddot{x}_k & \ddot{y}_k & \ddot{z}_k \end{bmatrix}^\top \tag{4}$$

Using a constant-acceleration motion model, the prediction step is given as:

$$\mathbf{x}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1}, \tag{5}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^\top + \mathbf{Q}_k, \tag{6}$$

where $\mathbf{F}$ is the 9×9 state transition matrix based on sampling interval Δt, and $\mathbf{Q}_k = \mathbf{G}\mathbf{Q}_0\mathbf{G}^\top$ is the process noise covariance. During the update step, when actual sensor measurements are incorporated, the measurement model can be defined as:

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_{k|k-1} + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k). \tag{7}$$

The Kalman gain and state update equations are following:

$$\mathbf{y}_k = \mathbf{z}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}, \tag{8}$$

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^\top + \mathbf{R}_k, \tag{9}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^\top\mathbf{S}_k^{-1}, \tag{10}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k\mathbf{y}_k, \tag{11}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}. \tag{12}$$

To address time-varying uncertainties in the dynamic environments, AKF adaptively updates measurement noise covariance $\mathbf{R}_k$ using the empirical innovation covariance over a sliding window of size N. Let $\mathbf{y}_i$ be the innovation vector at time step i over a fixed window of size N. The sample innovation covariance is:

$$\hat{\mathbf{C}} = \frac{1}{N-1} \sum_{i=k-N+1}^{k} \mathbf{y}_i\mathbf{y}_i^\top \tag{13}$$

We update the measurement noise covariance as:

$$\mathbf{R}_k = \hat{\mathbf{C}} - \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^\top \tag{14}$$

Additionally, the filter applies GDOP-based dynamic scaling to $\mathbf{R}_k$ to account for geometric uncertainty in UWB positioning. The scaled UWB measurement noise is computed as:

$$\mathbf{R}_{\text{UWB}}^{\text{scaled}} = \alpha(g)^2 \cdot \mathbf{R}_{\text{UWB}} \tag{15}$$

where the scaling factor $\alpha(g)$ is a function of the GDOP value g. The combined approach of using both residual-based

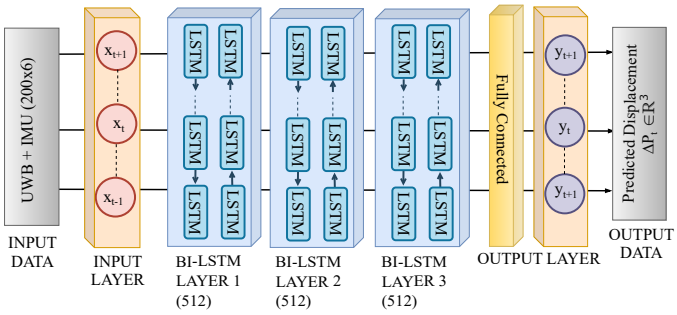


Figure 4. Architecture of the proposed Bi-LSTM fusion model with three hidden layers containing 512 neurons each. The model processes sequences of IMU and UWB data and predicts 3D displacement vectors for trajectory reconstruction.

adaptation and GDOP-based scaling makes the filter more reliable when sensor noise increases or anchor placement is poor. It is especially useful for the vertical Z-axis, where UWB measurements are often less accurate, because of the limited height variation in anchor positions.

IV. Sensor Fusion via Bi-LSTM Neural Network

A. Model Architecture

As an alternative approach to model-based filtering, we developed a learning-based sensor fusion technique based on a deep recurrent model using Bi-LSTM. The primary purpose of the model is to estimate the motion of the system by understanding the complex temporal relationships between IMU acceleration measurements and UWB trilaterated position estimates. In this fusion method, we adopted the frame-wise prediction strategy in which the model is trained to predict frame-wise displacement vectors. Here, a frame refers to a single-time instance incorporating input features that contain both IMU and UWB data. Each predicted displacement vector represents the relative positional change between two consecutive time steps, which can also be referred to as the delta position $\Delta\mathbf{p}_t$. Instead of directly predicting absolute positions between successive time steps, the proposed neural network focuses on learning the motion patterns specifically, how the system advances from one time step (or frame) to the next. The incremental formulation helps the system to generalize better in terms of understanding motion patterns and also aids in eradicating long-term drifts.

The model takes input samples in the form of a fixed-length sliding window of 200 consecutive time steps. For each time step, the input vector contains six features from the input data, which are three-axis accelerometer readings and also three-axis UWB positions. This combination forms an input sequence of $[200 \times 6]$ shape where temporal continuity and sensor fusion occur simultaneously. The input data is divided into three categories: 70% for training, 15% for validation, and 15% for testing. During the training phase, the model sequentially predicts the displacement vectors $\Delta\mathbf{p}_t \in \mathbb{R}^3$ from the input data at every time step in the sliding window, that is called the per-frame delta predictions. In the inference phase (testing), cumulative integration of these predicted displacement vectors generates a full 3D trajectory.

The bidirectional nature of the model allows processing of the input sequences in both forward and backward directions. The proposed model is composed of three stacked Bi-LSTM layers, each layer with a hidden dimension of 512 units per direction (forward and backward). To convert sequential features generated by stacked Bi-LSTM layers from the input data, the model uses a fully connected (dense) layer. The fully connected layer linearly maps the final hidden state at each time step to a 3D displacement vector. The bidirectional functionality helps the model to capture both past and future context for each timestep in the input window, rendering it particularly beneficial for modeling complex motion patterns in the sequential data. The detailed layer-wise architecture of the proposed Bi-LSTM network is depicted in Figure 4.

The training of the model is performed on the mini-batches of size 64 and optimized using adaptive moment estimation (Adam) optimizer with its standard learning rate of 0.001 to ensure stable and fast convergence. The loss function is based on a WMSE emphasizing prediction accuracy along the Z-axis. This weighting strategy is motivated by the notion of noise dominance across the vertical (Down) axis. The loss for one sample can be expressed as:

$$loss(x, y, z) = w_x(\hat{x} - x)^2 + w_y(\hat{y} - y)^2 + w_z(\hat{z} - z)^2 \quad (16)$$

where, $\hat{x}, \hat{y}, \hat{z}$ symbolizes the predicted delta positions, and x, y, z are the corresponding ground truth values. The weights w_x, w_y, w_z are adjustable according to the axis-specific importance. To formulate this loss over a batch of size N, the averaged WMSE $\mathcal{L}$ can be calculated as:

$$\mathcal{L} = \frac{1}{N}\sum_{n=1}^{N}\left[w_x(\hat{x}_n - x_n)^2 + w_y(\hat{y}_n - y_n)^2 + w_z(\hat{z}_n - z_n)^2\right] \quad (17)$$

The corresponding hyperparameters used during training are summarized in Table I. The model configuration was empirically optimized for the best performance. The reconstructed 3D trajectory from the Bi-LSTM model was obtained by integrating predicted position deltas over time and subsequently benchmarked against baseline methods and ground truth measurements.

V. Performance Evaluation

In the following, we evaluate the localization accuracy of the proposed Bi-LSTM sensor fusion model compared with classical approaches, including UWB-only positioning, IMU-only dead reckoning, and an AKF. Results are assessed using multiple metrics: root mean square error (RMSE), mean absolute error (MAE), maximum error, axis-wise decomposition, and statistical error distributions. Each evaluation is compared to the RTK GNSS ground truth. UWB-only serves as the baseline, IMU-only captures uncorrected inertial drift, AKF represents traditional model-based fusion, and Bi-LSTM embodies the learning-based method trained to reduce spatial

Table I
BI-LSTM MODEL HYPERPARAMETERS

Hyperparameter	Value	Description
Input Features	6	3-axis IMU + 3-axis UWB
Output Size	3	Δx, Δy, Δz prediction
Hidden Size	512	Per direction (Bi-LSTM)
Number of Layers	3	Stacked Bi-LSTM layers
Bidirectional	Yes	Forward and backward temporal modeling
Window Size	200	Length of input sequence
Batch Size	64	Selected via grid search
Learning Rate	1e-3	Selected via grid search
Optimizer	Adam	Adaptive gradient descent
Loss Function	Weighted MSE	Higher weight on Z-axis
Early Stopping	30 epochs	Patience based on val. loss
Data Split	70/15/15	Train/Validation/Test

localization error. Relevant configuration data is summarized in Table I.

A. Training Convergence and Stability

Figure 5 presents the training evolution of the proposed Bi-LSTM-based sensor fusion model. The top subplot shows the validation RMSE across training epochs, and the bottom subplot displays both the training and validation loss curves of WMSE on a logarithmic scale. The validation RMSE exhibits a rapid decline during the initial epochs, reflecting efficient early learning. Over time, minor fluctuations are observed in the later epochs, potentially caused by sensor noise and non-uniform characteristics in the validation data. Nevertheless, the overall trend remains stable with no signs of overfitting or divergence. The close alignment of the training and validation loss curves throughout the training further strengthens the model's generalization capability. The training phase is keeping track of the validation loss with an early stopping mechanism that stops training when no significant improvement occurs in the validation performance after a minimum of 30 epochs. Rather than training for all epochs, this method enhances computational efficiency by ceasing the training earlier, thereby avoiding overfitting and unnecessary performance degradation. In this way, the network produces per-

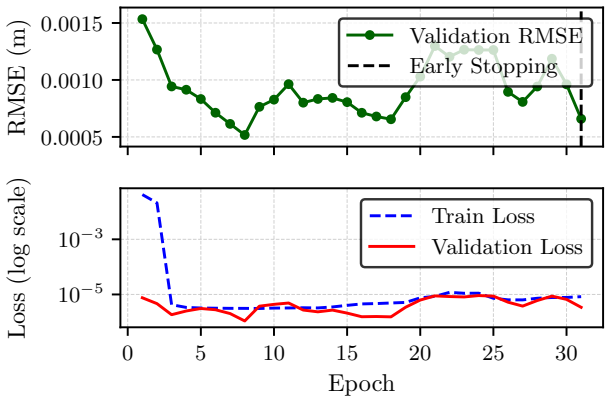


Figure 5. Bi-LSTM model training: validation RMSE (top), training and validation loss (bottom).

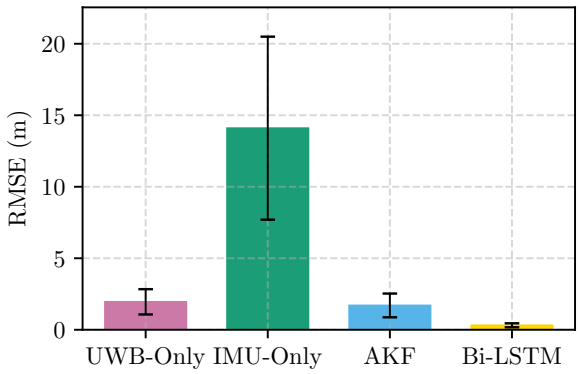


Figure 6. Average localization RMSE (Euclidean norm) across methods; error bars represent variance for each method.

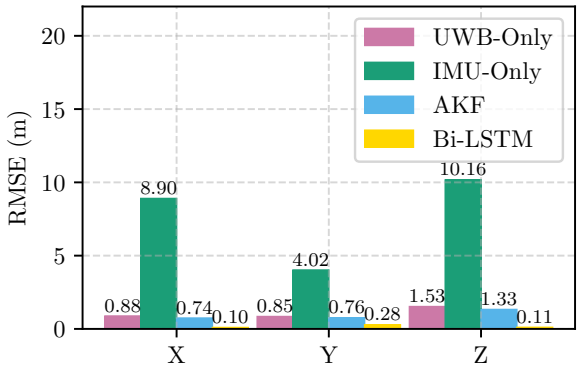


Figure 7. Average localization RMSE along X-, Y-, and Z-axes.

step delta displacement vectors $\Delta \mathbf{p}_t$, which are incrementally accumulated to reconstruct the complete 3D trajectory.

While we have optimized our model using WMSE loss function, however, RMSE validation is preferred for visualization and reporting as it offers a direct measure of spatial accuracy in physical units. The resulting curves demonstrate the model's generalization ability through stable convergence highlighting the robustness of the proposed model.

B. Overall Localization Accuracy

We first evaluate the accuracy of the location estimations with respect to ground truth using the Euclidean norm based RMSE (cf. Figure 6). As can be seen, the IMU-only approach leads to very high measurement errors. This effect is well-understood due to accumulating error over time. The UWB-only approach is better but still leads to high errors. The AKF-based fusion shows moderate reduction of RMSE. The proposed Bi-LSTM-based approach effectively decreases the RMSE to about 0.31 m, outperforming all other methods.

To further investigate positioning accuracy and directional performance, Figure 7 presents the RMSE for each spatial axis. In order to quantify the influence of noise, we particularly concentrated on the Z-axis since errors are more prominent in Down axis. For UWB, the calculated RMSE of 1.53 m is decreased by AKF to 1.33 m. However, our Bi-LSTM deep neural network achieves the best performance of 0.11 m, thus

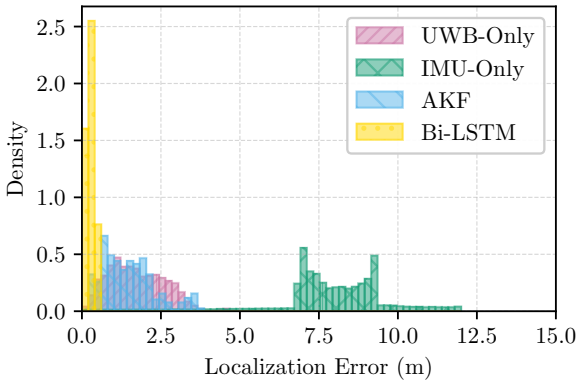


Figure 8. Histogram of localization errors.

enabling more precise elevation correction. This eventually proves that the trajectory estimation from the predicted delta positions has certainly closer alignment to the ground truth. Our Bi-LSTM approach helps to reduce both systematic drift in the IMU measurements and noisy fluctuations in UWB positions by learning context-aware temporal patterns.

C. Statistical Error Distribution

To further highlight the statistical error distribution, Figure 8 shows a histogram of localization errors. The IMU-only errors are the largest, even exceeding 10 m with a wide distribution due to accumulated integration drift in the measured data. The UWB-only errors extend beyond 4 m and concentrate around 2 m. The AKF reduces the error and has a slightly squeezed error distribution (but with peaks also reaching 4 m). In contrast, the Bi-LSTM model produces sharply concentrated errors with most errors remaining well below 0.5 m.

VI. Conclusion

In this paper, we presented a Bi-LSTM-based heterogeneous sensor fusion system for accurate 3D localization in dynamic and obstructed environments. The proposed model learns the spatiotemporal relationships between high-frequency IMU data and sparse UWB measurements to capture motion dynamics and handle sensor degradation or dropouts. The Bi-LSTM fusion model achieves significant improvements in localization accuracy – we achieved a 68 % reduction in RMSE on real-world field data recorded at the avalanche test site. We tuned the system to focus most notably on the vertical Z-axis by learning adaptive error patterns that account for modality-specific noise and temporal dynamics. In contrast to traditional filtering approaches with fixed noise assumptions, this data-driven method offers greater flexibility and resilience in diverse sensing conditions. These results highlight the effectiveness of deep recurrent fusion models in safety-critical applications. The proposed model is currently deployed for offline post-processing, where inference latency is not critical. Future work will explore real-time deployment and the scalability of the system to support collaborative operation across several sensing nodes in shared or distributed environments.

Acknowledgments

This work was supported by the project AvaRange funded by German Research Foundation (DFG), grant DR 639/22-1 and Austrian Science Fund (FWF), grant I 4274-N29.

References

[1] S. Wang and N. S. Ahmad, "A Comprehensive Review on Sensor Fusion Techniques for Localization of a Dynamic Target in GPS-Denied Environments," *IEEE Access*, pp. 2252–2285, Dec. 2024.

[2] M. Neuhauser, A. Köhler, R. Neurauter, M. S. Adams, and J.-T. Fischer, "Particle trajectories, velocities, accelerations and rotation rates in snow avalanches," *Annals of Glaciology*, pp. 1–18, Oct. 2023.

[3] F. Erlacher, B. Weber, J.-T. Fischer, and F. Dressler, "AvaRange - Using Sensor Network Ranging Techniques to Explore the Dynamics of Avalanches," in *12th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2016)*, Cortina d'Ampezzo, Italy: IEEE, Jan. 2016, pp. 120–123.

[4] J. Kuß, A. Köhler, M. Neuhauser, J.-T. Fischer, R. Neurauter, J. Gerstmayr, and F. Dressler, "A Measurement System for Distributed UWB-based Ranging and Localization in Snow Avalanches," in *29th ACM International Conference on Mobile Computing and Networking (MobiCom 2023), Poster Session*, Madrid, Spain: ACM, Oct. 2023.

[5] J. Sun, R. Gu, S. Li, S. Ma, H. Wang, Z. Li, and W. Feng, "HUID: DBN-based fingerprint localization and tracking system with hybrid UWB and IMU," *IEEE China Communications*, vol. 20, no. 2, pp. 139–154, Feb. 2023.

[6] A. Sesyuk, S. Ioannou, and M. Raspopoulos, "A Survey of 3D Indoor Localization Systems and Technologies," *MDPI Sensors, Feature Papers in Communications Section*, vol. 22, Dec. 2022.

[7] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Non-Line-of-Sight Identification and Mitigation Using Received Signal Strength," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1689–1702, Mar. 2015.

[8] V. Barral Vales, C. J. Escudero, J. A. García-Naya, and R. Maneiro-Catoira, "NLOS Identification and Mitigation Using Low-Cost UWB Devices," *MDPI Sensors*, vol. 19, no. 16, Aug. 2019.

[9] L. Zhang, C. Wang, M. Ma, and D. Zhang, "WiDIGR: Direction-Independent Gait Recognition System Using Commercial Wi-Fi Devices," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1178–1191, Feb. 2020.

[10] J. Zhang, J. Salmi, and E.-S. Lohan, "Analysis of Kurtosis-Based LOS/N-LOS Identification Using Indoor MIMO Channel Measurement," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2871–2874, Jul. 2013.

[11] Y. Li, Y. Zhuang, P. Zhang, H. Lan, X. Niu, and N. El-Sheimy, "An improved inertial/wifi/magnetic fusion structure for indoor navigation," *Information Fusion*, pp. 101–119, Mar. 2017.

[12] Y. Li, Z. He, Z. Gao, Y. Zhuang, C. Shi, and N. El-Sheimy, "Toward Robust Crowdsourcing-Based Localization: A Fingerprinting Accuracy Indicator Enhanced Wireless/Magnetic/Inertial Integration Approach," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3585–3600, 2019.

[13] Y. Zhuang, J. Yang, L. Qi, Y. Li, Y. Cao, and N. El-Sheimy, "A Pervasive Integration Platform of Low-Cost MEMS Sensors and Wireless Signals for Indoor Localization," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4616–4631, Dec. 2018.

[14] C. Chen and X. Pan, "Deep Learning for Inertial Positioning: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 10 506–10 523, Sep. 2024.

[15] M. Narasimhappa, A. D. Mahindrakar, V. C. Guizilini, M. H. Terra, and S. L. Sabat, "MEMS-Based IMU Drift Minimization: Sage Husa Adaptive Robust Kalman Filtering," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 250–260, Jan. 2020.

[16] D. Feng, C. Wang, C. He, Y. Zhuang, and X.-G. Xia, "Kalman-Filter-Based Integration of IMU and UWB for High-Accuracy Indoor Positioning and Navigation," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133–3146, Apr. 2020.

[17] A. Benini, A. Mancini, A. Marinelli, and S. Longhi, "A Biased Extended Kalman Filter for Indoor Localization of a Mobile Agent using Low-Cost IMU and UWB Wireless Sensor Network," *Elsevier IFAC Proceedings Volumes*, vol. 45, no. 22, pp. 735–740, 2012.

[18] Z. Liu, Y. Li, Y. Wu, and S. He, "Formation control of nonholonomic unmanned ground vehicles via unscented Kalman filter-based sensor fusion approach," *ISA Transactions*, vol. 125, pp. 60–71, Jun. 2022.

[19] Q. Tian, K. I.-K. Wang, and Z. Salcic, "A Resetting Approach for INS and UWB Sensor Fusion Using Particle Filter for Pedestrian Tracking," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5914–5921, Aug. 2020.

[20] H. A. Hashim, L. J. Brown, and K. McIsaac, "Nonlinear Stochastic Attitude Filters on the Special Orthogonal Group 3: Ito and Stratonovich," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 9, pp. 1853–1865, Sep. 2019.

[21] C. Jiang, J. Shen, S. Chen, Y. Chen, D. Liu, and Y. Bo, "UWB NLOS/LOS Classification Using Deep Learning Method," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2226–2230, Oct. 2020.

[22] Z. Zhaoxia, X. Zhongwei, and J. Xia, "Deep learning optimization positioning algorithm based on UWB/IMU fusion in complex indoor environments," *Elsevier Physical Communication*, vol. 71, Aug. 2025.

[23] H. Liu, K. Li, Q. Fu, and L. Yuan, "Research on Integrated Navigation Algorithm Based on Radial Basis Function Neural Network," *Journal of Physics: Conference Series*, pp. 012–031, Jun. 2021.

[24] L. Cong, S. Yue, H. Qin, L. Bin, and J. Yao, "Implementation of a MEMS-Based GNSS/INS Integrated Scheme Using Supported Vector Machine for Land Vehicle Navigation," *IEEE Sensors Journal*, vol. 20, no. 23, pp. 14 423–14 435, Dec. 2020.

[25] W. Fang, J. Jiang, S. Lu, Y. Gong, Y. Tao, Y. Tang, P. Yan, H. Luo, and L. Bin, "A LSTM Algorithm Estimating Pseudo Measurements for Aiding INS during GNSS Signal Outages," *Remote Sensing, High-precision GNSS: Methods, Open Problems and Geoscience Applications*, vol. 12, 2020.

[26] Y. Zhang, "A Fusion Methodology to Bridge GPS Outages for INS/GPS Integrated Navigation System," *IEEE Access*, pp. 61 296–61 306, Apr. 2019.

[27] D. Li, Y. Wu, and J. Zhao, "Novel Hybrid Algorithm of Improved CKF and GRU for GPS/INS," *IEEE Access*, pp. 202 836–202 847, Nov. 2020.

[28] S. Chen, M. Xin, F. Yang, X. Zhang, J. Liu, G. Ren, and S. Kong, "Error Compensation Method of GNSS/INS Integrated Navigation System Based on AT-LSTM During GNSS Outages," *IEEE Sensors Journal*, vol. 24, no. 12, pp. 20 188–20 199, Jun. 2024.

[29] G. Feng, C. Shen, C. Long, and F. Dong, "GDOP index in UWB indoor location system experiment," in *IEEE SENSORS*, Busan, South Korea, Nov. 2015, pp. 1–4.

[30] A. Poulose, Z. Emersic, O. Steven Eyobu, and D. S. Han, "An Accurate Indoor User Position Estimator For Multiple Anchor UWB Localization," in *11th International Conference on Information and Communication Technology Convergence (ICTC 2020)*, Jeju Island, South Korea: IEEE, Oct. 2020, pp. 478–482.

[31] B. Van Herbruggen, D. Van Leemput, J. Van Landschoot, and E. De Poorter, "Real-Time Anchor Node Selection for Two-Way-Ranging (TWR) Ultra-Wideband (UWB) Indoor Positioning Systems," *IEEE Sensors Letters*, vol. 8, no. 3, pp. 1–4, Feb. 2024.

[32] A. Beauvisage, K. Ahiska, and N. Aouf, "Robust Multispectral Visual-Inertial Navigation With Visual Odometry Failure Recovery," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9089–9101, Jul. 2022.

[33] J. Cheng, L. Yang, Y. Li, and W. Zhang, "Seamless outdoor/indoor navigation with WIFI/GPS aided low cost Inertial Navigation System," *Elsevier Physical Communication*, pp. 31–43, Jan. 2014.

[34] Q. Zou, Q. Sun, L. Chen, B. Nie, and Q. Li, "A Comparative Analysis of LiDAR SLAM-Based Indoor Navigation for Autonomous Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6907–6921, Jul. 2022.

[35] S. Guan and X. Luo, "Fusing Ultra-wideband Range Measurements with IMU for Mobile Robot Localization," in *11th International Conference on Intelligent Control and Information Processing (ICICIP 2021)*, Dali, China: IEEE, Dec. 2021, pp. 107–111.

[36] P. Ji, Z. Duan, and W. Xu, "A Combined UWB/IMU Localization Method with Improved CKF," *MDPI Sensors*, vol. 24, May 2024.

[37] M. Dong, Y. Qi, X. Wang, and Y. Liu, "A Non-Line-of-Sight Mitigation Method for Indoor Ultra-Wideband Localization With Multiple Walls," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 7, pp. 8183–8195, Jul. 2023.

[38] A. Bonci, E. Caizer, M. C. Giannini, F. Giuggioloni, and M. Prist, "Ultra Wide Band communication for condition-based monitoring, a bridge between edge and cloud computing," *Procedia Computer Science*, vol. 217, pp. 1670–1677, 2023.

[39] L. Yao, Y.-W. A. Wu, L. Yao, and Z. Z. Liao, "An integrated IMU and UWB sensor based indoor positioning system," in *7th IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN 2017)*, Sapporo, Japan: IEEE, Sep. 2017, pp. 1–8.

[40] X. Li, J. Ye, Z. Zhang, F. Liang, W. Xia, and B. Wang, "A Novel Low-Cost UWB/IMU Positioning Method With the Robust Unscented Kalman Filter Based on Maximum Correntropy," *IEEE Sensors Journal*, vol. 24, no. 18, pp. 29 219–29 231, Sep. 2024.

[41] Y. Lee and D. Lim, "Vision/UWB/IMU sensor fusion based localization using an extended Kalman filter," in *IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE 2019)*, Yunlin, Taiwan: IEEE, Oct. 2019, pp. 401–403.

[42] Y. Kuang, T. Hu, M. Ouyang, Y. Yang, and X. Zhang, "Tightly Coupled LIDAR/IMU/UWB Fusion via Resilient Factor Graph for Quadruped Robot Positioning," *Remote Sensing*, vol. 16, Nov. 2024.

[43] R. Neurauter, A. Holzinger, M. Neuhauser, J.-T. Fischer, and J. Gerstmayr, "Motion Reconstruction of Fast-rotating Rigid Bodies," *Journal of Computational and Nonlinear Dynamics*, vol. 19, no. 1, p. 011 005, Jan. 2024.

[44] M. Tommingas, T. Laadung, S. Varbla, I. Müürsepp, and M. Mahtab Alam, "UWB and GNSS Sensor Fusion Using ML-Based Positioning Uncertainty Estimation," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 2177–2189, 2025.

[45] Z. Zhi, D. Liu, and L. Liu, "A performance compensation method for GPS/INS integrated navigation system based on CNNâ€"LSTM during GPS outages," *Measurement*, vol. 188, Jan. 2022.

[46] Y. Song and L.-T. Hsu, "Tightly coupled integrated navigation system via factor graph for UAV indoor localization," *Aerospace Science and Technology*, vol. 108, Jan. 2021.

[47] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, and F. Fioranelli, "Bi-LSTM Network for Multimodal Continuous Human Activity Recognition and Fall Detection," *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1191–1201, Feb. 2020.

[48] J. Kuß, A. Köhler, M. Neuhauser, R. Neurauter, J. Gerstmayr, J.-T. Fischer, and F. Dressler, "Distributed UWB-based Ranging for Particle Tracking in Avalanches," in *19th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2024)*, Chamonix, France: IEEE, Jan. 2024, pp. 125–132.

[49] R. Neurauter, P. Hergel, and J. Gerstmayr, "Evaluation of Inertial Measurement Units for Short Time Motion Tracking," in *17th International Conference on Multibody Systems, Nonlinear Dynamics, and Control (MSNDC 2021)*, Virtual Conference: ASME, Aug. 2021.

[50] R. Neurauter and J. Gerstmayr, "A novel motion-reconstruction method for inertial sensors with constraints," *Multibody System Dynamics*, vol. 57, no. 2, pp. 181–209, Dec. 2022.

[51] R. Winkler, M. Neuhauser, R. Neurauter, F. Erlacher, W. Steinkogler, and J.-T. Fischer, "Particle tracking in snow avalanches with in situ calibrated inertial measurement units," *Annals of Glaciology*, vol. 65, e14, Jan. 2024.

[52] D. Neirynck, E. Luk, and M. McLaughlin, "An alternative double-sided two-way ranging method," in *13th Workshop on Positioning, Navigation and Communications (WPNC 2016)*, Bremen, Germany: IEEE, Oct. 2016.