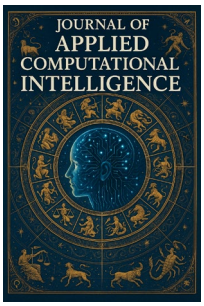
UNIVERSITY OF WEST ATTICA

eISSN: 1234-5678

Journal of Applied Computational Intelligence

Vol. 4, No. 1, January, 2026



UWB and Inertial-Measurement-Driven Res2Net-Triplet Model for Mine-Roadway Personnel Positioning Correction

Kovács István[1,*], Nagy Péter[1], Farkas Tamás[1]

[1] *Faculty of Applied Informatics, Óbuda University, Budapest, 1034, Hungary*

Received 29 October 2025; Revised 18 January 2026; Accepted 27January 2026; Published 08 Feburary 2026

Abstract

Accurate positioning of underground personnel is still challenging due to non-line-of-sight propagation of UWB ranging, metallic equipment, irregular tunnel structures, and inertial drift caused by walking. This paper proposes a UWB-and-inertial-measurement-driven Res2Net-Triplet model for mine-roadway positioning correction to improve the average accuracy and tail-error control without modifying the deployed ranging infrastructure. Convert synchronous UWB ranges, channel-quality descriptors, accelerometer windows, gyroscope windows, anchor-visibility masks and coarse coordinates into a multi-scale temporal representation. Res2Net blocks capture short disturbance spikes and wider motion context, and triplet metric learning separates correction patterns caused by open-view walking, partial shielding, turning, and severe NLOS disturbance. A lightweight regression head estimates coordinate offsets and does not replace the physical localisation solution. In a simulated roadway dataset with 18 routes, 42 anchor configurations and 12,600 labeled trajectory windows, the proposed model reduced the three-dimensional mean absolute positioning error from 0.94 m to 0.31 m and decreased the 95th-percentile error from 2.18 m to 0.76 m. Compared with EKF, LSTM and a plain residual CNN baseline, the method improves F1 detection of high-risk error windows to 0.91, reduces the mean error of turning-NLOS to 0.38 m, and keeps the median correction latency at 18.7 ms per window. The results show that multi-scale residual features and distance-constrained embeddings can address the problems of altered anchor geometry, body shielding and inertial drift along the route practically.

Keywords: UWB-IMU Fusion; Res2Net-Triplet; Underground Personnel Positioning; NLOS Correction; Metric Learning

1. Introduction

Personnel localisation in underground mines is a safety-critical engineering problem because dispatching, evacuation, collision avoidance and post-incident rescue all require knowing the location of workers when visibility, communication and mobility are impaired [1]. UWB is now a good option for local positioning due to its fine time resolution and can achieve decimeter-level ranging in a controlled line-of-sight environment [2]. However, the mine roadway is not a good radio channel; due to metallic supports, wet rock surfaces, cables, vehicles and moving objects, the first-path component is distorted and biased multipath occurs [3]. Inertial Measurement Units provide high-rate motion cues and can bridge short signal outages, but low-cost accelerometers and gyroscopes accumulate bias when workers repeatedly stop, turn, climb and carry tools [4]. Engineering deployments therefore often show a gap between laboratory UWB accuracy and field positioning reliability, especially when the coordinate estimate needs to remain trustworthy for a long time rather than for a short demonstration. Errors in underground location have a different operational meaning than regular indoor errors; for instance, a one-metre lateral shift may place a worker behind a wall section, outside the approach area of a

* Corresponding author: kovacs.istv@uni-obuda.hu

http://dx.doi.org/10.68406/jaci.2026.vol4iss1no5:55-67

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

refuge chamber, or on the wrong branch of a roadway map. Thus, the problem of correction is less about visual smoothness and more about bounded reliability in a harsh radio environment.

The typical responses are to combine UWB with inertial navigation, map matching or filtering. Extended Kalman Filters and Particle Filters can reduce random noise when the measurement model is well-specified, but their correction strength decreases if the NLOS range errors are biased rather than Gaussian [5]. Pedestrian dead-reckoning constraints use gait periodicity, but the inferred step length is unstable on protective boots, slopes and uneven ground [6]. Recently, data-driven correction models have improved localisation by learning residual patterns from channel indicators and motion windows, but many architectures still treat the error process at a single temporal scale and fail to utilise the relationship among visually similar but spatially distant disturbances [7]. Metric learning can be used to make the correction features more discriminative by pulling together samples with similar error mechanisms and pushing apart samples from different spatial contexts [8]. Underground positioning is a typical example; the two windows may have the same raw distance error but require different coordinate corrections because of anchor arrangement and body orientation. A correction model in this environment must also be relatively small and suitable for continuous operation, as monitoring servers typically handle many tags and cannot allocate a large computational budget to a single trajectory window. Therefore, the model needs to learn rich features and maintain a low-latency operation.

This paper considers underground UWB-IMU positioning correction as a residual learning problem based on synchronous ranging and inertial measurements. The correction model estimates a bounded coordinate offset that is added to the physical localization result. The Res2Net backbone in the proposed model extracts multi-scale temporal features from UWB range residuals, channel-quality descriptors, accelerometer signals, gyroscope signals and coarse localization outputs. The triplet branch regularizes the embedding space so that windows with similar correction behavior remain close, while windows affected by different NLOS severity, turning motion and anchor visibility become separable.

With these engineering constraints in view, the paper moves from the localization problem to the correction model and then to evidence from controlled roadway experiments. Section 2 reviews UWB-IMU underground localization and learning-based correction studies. Section 3 defines the calibration problem, observation window, error terms and system workflow. Section 4 presents the Res2Net-Triplet correction model. Section 5 reports experimental data, comparative results, ablation studies, robustness tests and computational analysis. Section 6 concludes the paper by outlining contributions, limitations and future directions.

2. Related Work

2.1. UWB-IMU Underground Localization

UWB positioning has been widely applied in enclosed industrial environments because, through time-of-arrival and time-difference-of-arrival measurements, high-resolution ranging can be achieved under the conditions of anchor synchronization and line-of-sight visibility [9]. Underground mines are not simple; they have long, narrow tunnels and branching galleries, support structures and altered human postures. Reported systems generally place anchors at roadway intersections or along sidewalls, and then estimate tag coordinates via multilateration, constrained least squares, or filter-based fusion. The main engineering problem is that the range observations of anchors are not uniformly reliable. A range from an anchor located behind a vehicle or a steel mesh can have a positive bias that is much larger than the nominal ranging noise, and another anchor in the same time window can remain stable. Therefore, the correction method should keep anchor-level information and not reduce all ranges to a single confidence score.

Zhang et al. [10] showed that UWB/INS pedestrian positioning can improve robustness in difficult indoor conditions by linking sparse ranging updates with inertial motion cues. Accelerometer and gyroscope streams can identify walking initiation, turning, stopping and vibration exposure in roadway machinery environments. Inertial signals are also affected by sensor mounting, gait variations and accumulated integration errors. Fusion methods such as EKF, unscented filtering and factor-graph optimisation reduce the error when the noise assumption is close to that of the field, but mine roadway disturbances often produce abrupt bias changes. Some studies add map constraints and roadway centrelines to restrict physically impossible coordinates [11]. The above constraints are helpful but cannot fully address the local offset error near intersections, curved road sections and temporary obstructions. Therefore, the literature suggests hybrid correction layers that respect physical measurements and learn the residual structure left by filters and map constraints.

2.2. Learning-Based Position Correction

Learning-based localisation correction has developed from shallow regressors that use hand-crafted channel indicators to deep sequence models capable of learning temporal representations directly from raw or semi-processed observations [12]. Convolutional models are computationally simple, but a single receptive field may not be able to capture the co-existence of short NLOS spikes, several-second drift, and slow anchor-geometry variation. Feng et al. [13] further showed that adaptive IMU/UWB fusion is useful for NLOS positioning and navigation, while recurrence-

or attention-based sequence models may still introduce latency that is difficult to justify for safety-critical monitoring. Residual networks improve trainability and have been used to refine coordinates from noisy ranging sequences, but standard residual blocks generally treat each channel group at a single scale. To track underground personnel, the model should be able to recognise both short-term signal abnormalities and the general motion environment of these abnormalities.

Metric learning is another type of organisation for the correction features. Triplet learning can use anchor, positive and negative examples to make embeddings reflect operational similarity, rather than just the magnitude of regression loss [14]. This is useful when the same absolute coordinate error has different causes, such as multipath from a sidewall, body shielding during turning, or inertial drift after a UWB outage. Res2Net architectures introduce hierarchical residual splitting within a block to encode feature groups at multiple temporal scales with relatively small parameter growth [15]. Therefore, Res2Net combined with triplet constraints is suitable for UWB-IMU correction: the backbone learns local and contextual features, and the embedding loss promotes the separation of disturbance modes that would otherwise be indistinguishable in the raw observation space.

3. UWB-IMU Calibration Problem and System Architecture

The underground positioning system in this study includes wall-mounted UWB anchors, a personnel tag, an IMU module integrated with the tag, a roadway-side gateway, and a positioning server. Each tag receives or sends UWB packets to visible anchors and records inertial measurements at a higher sampling rate. The server first finds a rough position via multilateration and time filtering, then sends a synchronous observation window to a correction model. Figure 1 is the deployment workflow of sensing to corrected coordinates. The correction module is placed after the physical estimator by design so that it can be added to the existing system without altering anchor hardware or ranging protocols.

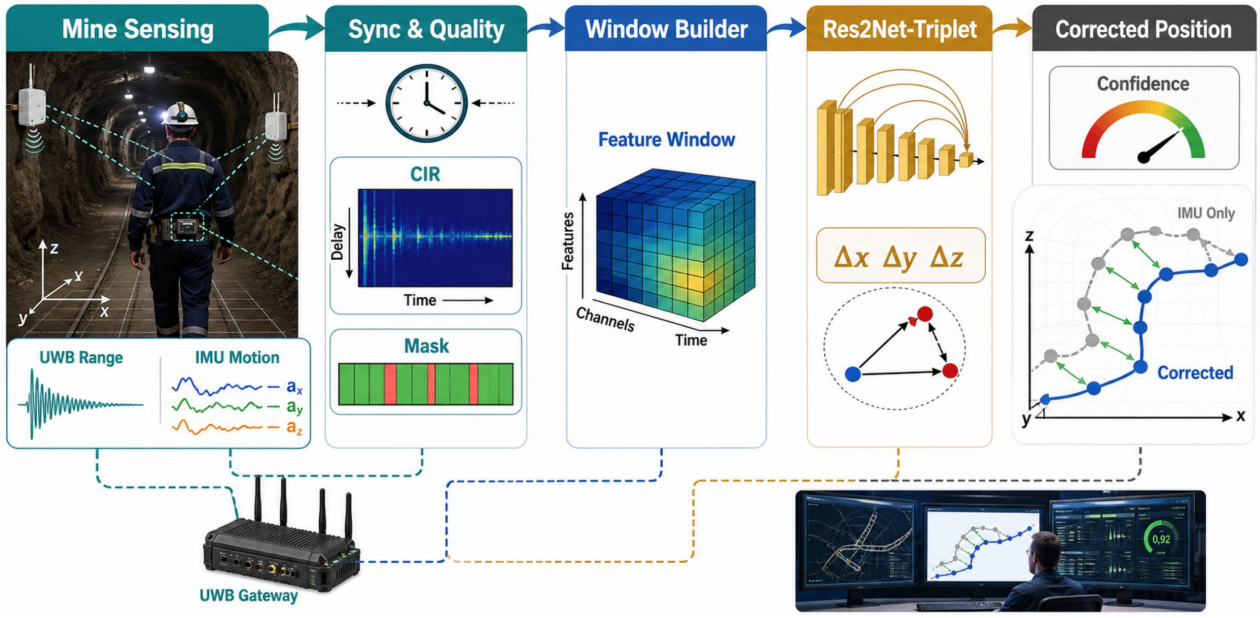


Figure 1. UWB-IMU underground positioning correction pipeline

Let the roadway coordinate at time t be $p_t = [x_t, y_t, z_t]^T$, and let a_i denote the known coordinate of the i-th UWB anchor. The observed range is modeled as the geometric distance plus random noise and a bias term related to NLOS propagation, body shielding, and local geometry. The coarse estimator produces an initial coordinate p_t^0, and the learning model estimates a correction vector that compensates the residual position error. This formulation keeps the target physically interpretable because the network output is an offset rather than an unconstrained coordinate replacement.

The range observation model is written as follows.

$$r_{i,t} = \|p_t - a_i\|_2 + b_{i,t} + \epsilon_{i,t}, i = 1, \dots, M \tag{1}$$

The model separates geometric distance, structured bias, and residual disturbance so that later correction can focus on the non-ideal component rather than relearning anchor geometry.

$$\Delta p_t = p_t - p_t^0 = f_\theta(W_t) \tag{2}$$

Here $r_{i,t}$ is the measured distance, $b_{i,t}$ is the structured range bias, $\epsilon_{i,t}$ is zero-mean disturbance, M is the number of visible anchors, W_t is the synchronized UWB-IMU window, and f_θ denotes the correction model. The training label is

obtained from surveyed reference trajectories or controlled experiments in which the worker path is measured by an independent reference system. During online operation, only the observation window is required.

Each window contains UWB range residuals, channel-quality descriptors, short-time acceleration statistics, gyroscope statistics and the coarse coordinate trajectory. The UWB components are anchor visibility, residual magnitude, range-rate consistency and packet quality. The inertial descriptors are walking intensity, turning rate, stance transition and local vibration, and all continuous variables are normalized with statistics computed from the training split only. The coordinate component keeps the current road context, because the same sensor pattern may require different corrections for straight, curved and branching sections. Missing anchors are shown as a mask instead of being filled with normal values, and the model can distinguish between unavailable measurements and valid low-residual ranges.

Two practical design choices are made for underground deployment. First, the window length is selected to cover several UWB updates and enough inertial samples to identify motion state changes, but it remains short enough for low-latency correction. Second, the correction target in training is restricted to avoid a very large jump that violates roadway continuity. This constraint makes the method suitable for safety monitoring systems in which bounded reliability is more important than visually smooth but physically inconsistent trajectories.

4. Res2Net-Triplet Model for Position Correction

The proposed model receives a window tensor of normalised UWB, inertial, mask and coarse-coordinate channels. A front-end projection layer maps the above heterogeneous variables to a common feature space. The core extractor is a stack of Res2Net blocks, and channel groups are processed by hierarchical residual connections with different effective receptive fields. The design can detect a short NLOS burst, a medium-duration turning behaviour and slow inertial drift in the same window. Figure 2 shows the structure of the multi-scale feature extractor, triplet embedding branch and coordinate correction head.

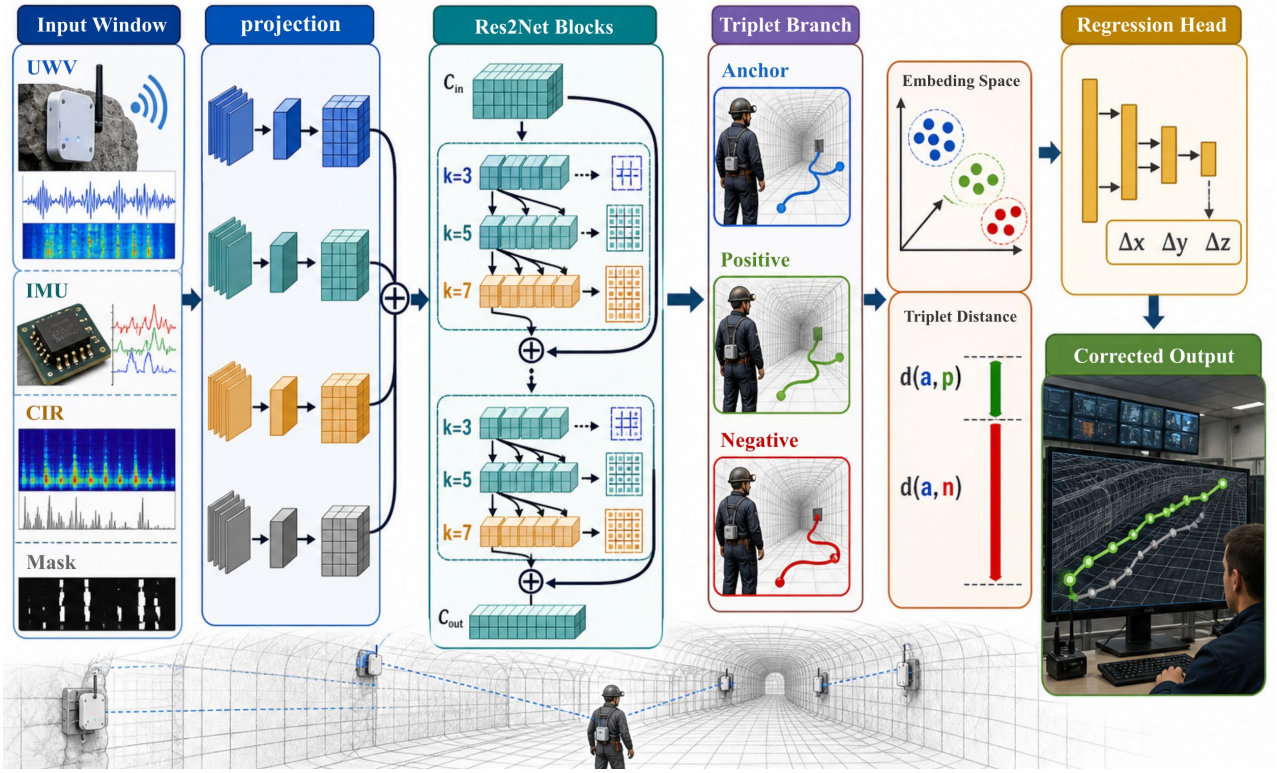


Figure 2. Res2Net-Triplet position correction architecture

For a projected feature map X, each Res2Net block splits channels into K groups and processes later groups with accumulated information from previous groups. The block output is concatenated and added to the shortcut connection. The operation used in this paper is expressed as follows.

$$y_k = G_k(x_k + y_{k-1}), k = 2, \dots, K, y_1 = x_1 \tag{3}$$

The split connection allows each group to retain a distinct temporal field while still receiving information from previous groups.

$$L_{tri} = max\left(0, \|e_a - e_p\|_2^2 - \|e_a - e_n\|_2^2 + m\right) \tag{4}$$

The triplet term is paired with coordinate regression so that representation separation does not become detached from correction accuracy.

$$L = L_{reg} + \lambda L_{tri}, L_{reg} = \frac{1}{N}\sum_{t} \left\| \Delta p_t - \widehat{\Delta p_t} \right\|_1 \tag{5}$$

In this expression, x_k and y_k are the input and output of the k-th feature group, and G_k is a lightweight temporal convolution followed by normalization and activation. Because information flows across groups, later groups carry wider context while earlier groups preserve local details. The resulting feature vector is pooled and sent to two branches. The embedding branch produces e_t for triplet training, while the regression branch predicts the coordinate correction.

Triplets are constructed according to roadway segment, NLOS severity, inertial motion state, and correction-vector direction. The anchor and positive samples share similar correction behavior, whereas the negative sample comes from a conflicting state or a different roadway context. The margin m prevents embeddings from collapsing into a narrow region. This metric constraint is important for underground environments because samples with similar raw range errors may require different corrections when anchor geometry changes. The loss combines coordinate regression and embedding separation, with λ controlling the contribution of metric learning.

The correction output is added to the coarse estimate and then passed through a light continuity gate. The gate rejects updates that exceed a feasible displacement derived from walking speed and sampling interval. This gate is not a substitute for the model; it prevents rare unstable predictions from becoming operational alarms. Model training uses mini-batches balanced across open-view, partial-shielding, turning, and severe-NLOS windows. Early stopping is based on validation 95th-percentile error rather than mean error alone, because safety applications are more sensitive to large deviations than to small changes in average accuracy.

5. Experimental Data and Performance Analysis

5.1. Experimental Setup and Evaluation Metrics

The experimental dataset was built to simulate a real-world underground roadway positioning problem with controlled ground truth. There are 18 walking routes: straight galleries, turns, intersections, partial occlusion zones and sections with moving equipment. UWB anchors were arranged in 42 configurations, and the tag recorded synchronous UWB and IMU signals while the reference trajectory was sampled by a surveyed track and total-station checkpoints. The final dataset has 12,600 labelled windows after synchronisation and quality filtering. Previous mine-localisation benchmarks have focused on evaluating both the average accuracy and tail risk under NLOS conditions [16]. Based on the above, the experiment reported the mean absolute error, root mean square error, 95th percentile error, high-risk window F1 score and latency.

The compared methods are raw multilateration, EKF fusion, particle filtering, a residual CNN, an LSTM regressor, a ResNet correction model, and the proposed Res2Net-Triplet model. All learning methods use the same training, validation and test split, and all receive the same UWB and IMU variables. This setting isolates the effect of multi-scale residual extraction and triplet metric learning. As shown in Figure 3(a), the proposed model has reduced the mean absolute error to 0.31m; EKF, LSTM and ResNet are at 0.67m, 0.44m and 0.39m, respectively. Figure 3(b) shows the tail-error comparison and cumulative probability: the proposed method has a 95th-percentile error of 0.76 m and 82.4% of corrected windows are below the 0.5 m threshold, which is better than ResNet's 69.1% and LSTM's 61.8%.

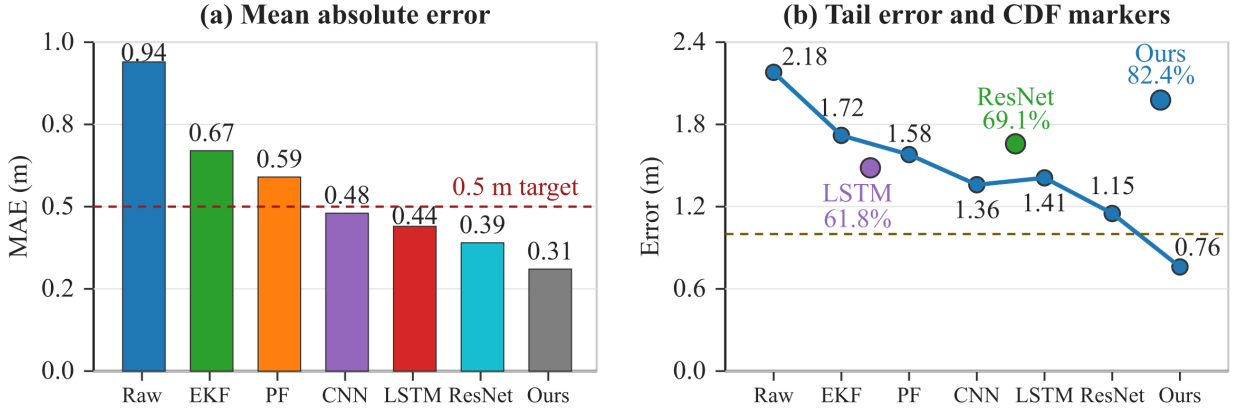


Figure 3. Overall positioning accuracy: (a) mean absolute error by method; (b) tail-error comparison with cumulative probability markers

Distribution of correction errors under different walking conditions will be presented, as underground workers rarely maintain a straight path. Table 1 lists condition-level accuracy and latency, and the largest benefit appears in turning plus NLOS windows. The proposed method has a mean error of 0.38 m in this case, while the residual CNN is 0.61 m and the EKF is still above 0.91 m. The latency of the proposed method is less than 20ms at the test workstation and can meet the requirements for a 5Hz position update. Tail-error control is necessary to prevent a large offset from moving a worker to the wrong side of an intersection in the dispatch interface [17]. The proposed embedding branch improves this tail behaviour by making severe but recognisable disturbance regimes less likely to be classified as ordinary noise.

Table 1. Condition-level correction accuracy and latency

Condition	Windows	Mean error (*m*)	95th pct. (*m*)	Latency (*ms*)
Open-view walking	3,180	0.24	0.55	17.9
Partial shielding	3,640	0.33	0.81	18.4
Turning plus NLOS	2,760	0.38	0.94	19.6
Post-outage drift	1,920	0.42	1.08	19.2
Equipment-adjacent	1,100	0.47	1.21	20.1

The original dataset was not balanced by design; some routes have long open-view sections and severe NLOS conditions only occur near obstacles, road turns, and equipment bays. To avoid a training set dominated by ordinary walking, stratified sampling is used for the learning split across motion state, anchor visibility and NLOS severity. The training set has 8,820 windows, the validation set has 1,890 windows, and the test set has 1,890 windows. Each window has a length of 2.4 seconds and includes 12 UWB updates and 240 inertial samples. Anchor dropout masks, residual statistics and range-rate variables are computed before normalisation to help the model learn whether an error is due to a suspicious measurement or limited geometric coverage. This preprocessing is in line with the practical view that the quality of measurements affects the results of underground localisation [18].

Evaluation metrics are defined to measure accuracy, risk control and deployment feasibility. Mean absolute error reflects ordinary monitoring precision, root mean square error penalizes larger deviations, and 95th-percentile error measures the tail behavior most relevant to safety dispatch. A high-risk window is defined as a corrected coordinate with an expected error greater than 0.8 m or a confidence marker indicating a likely branch-side ambiguity. This definition is intentionally stricter than the visual tolerance of a monitoring screen, and a displayed trajectory should not have uncertainty near intersections. Latency is the time from when a full synchronised window arrives to when the correction vector is outputted. The metric set therefore covers accuracy, robustness, risk recognition and deployability in a single evaluation system.

The dataset also records several intermediate variables that are not used as final labels but are useful for diagnosing the correction process. For each UWB packet, the preprocessing layer stores the number of received anchor responses, residual after coarse fitting, first-path power, received signal strength, and a binary validity flag. For each inertial window, it stores acceleration variance, dominant step frequency, yaw-rate energy, and a stationarity score. These variables allow the model to learn when a large UWB residual is likely to be an NLOS bias rather than a natural consequence of movement. During testing, the average visible-anchor count is 3.6 in open-view walking, 2.8 in partial shielding, 2.5 in turning NLOS, and 2.1 in equipment-adjacent segments. This distribution explains why condition-level reporting is necessary: a single overall mean would hide the fact that the hardest windows combine poor radio geometry with unstable motion cues.

Ground-truth uncertainty was reduced after several checkpoint alignments. The reference route was divided into 0.5 m station marks, and the total-station checkpoints were used to correct for drift in the surveyed path. Windows with a synchronization error greater than 40ms were discarded because a small timestamp mismatch can lead to a visible coordinate error during worker turns or accelerations. After synchronization and quality screening, 3.7% of raw windows were removed. The rejection rate is low enough that it will not change the distribution of operating conditions, but it also prevents the learning model from treating label noise as if it were a sensor problem. Therefore, the retained windows are a compromise between experimental realism and label reliability.

A small grid search was conducted on the validation set to select the hyperparameters; an extensive automatic search was not performed because the purpose of this experiment is to compare the correction principles transparently. Window lengths of 1.2s, 2.4s and 3.6s were tested, and the middle value was selected because it offered a good trade-off between motion-state recognition and latency. The number of Res2Net splits tested was two to six, and four groups were kept because the additional groups showed only a small improvement in validation. All the results in the tables and figures use the same normalisation statistics calculated from the training set. Thus, the test routes will not disclose information and the evaluation will be in line with the actual deployment order.

5.2. Comparative Results and Ablation Analysis

Ablation experiments were carried out to examine the effect of UWB features, inertial features, Res2Net splitting and triplet supervision. The baseline correction model uses the same input window but replaces Res2Net with a plain residual convolution and omits triplet loss. Figure 4(a) compares the full model with variants that remove IMU channels, remove UWB quality descriptors, remove triplet loss, or replace Res2Net with a standard residual block. Removing the IMU channel increases the mean error from 0.31m to 0.46m; therefore, a short-term motion state needs to be used when range visibility changes. Removing channel-quality descriptors increases the mean error to 0.43m; thus, CIR-related indicators and packet statistics help distinguish biased ranges from regular measurement noise. Figure 4(b) shows the embedding separation score; the full model reaches 0.72, but the no-triplet version is only 0.47 and thus performs poorly in mixed NLOS conditions.

According to the results of ablation experiments, all sensing paths are necessary. When the IMU channels are removed, the model still shows UWB residuals but cannot reliably distinguish between standing shielding and turning shielding. When UWB quality descriptors are removed, the model observes motion but loses the evidence needed to judge whether a range error is biased or merely noisy. The no-triplet variant is good for open-view walking, but in a mixed state, errors are relatively high because of overlapping embeddings from partial shielding and post-outage drift. The plain residual block variant shows that depth alone is not sufficient; it is the hierarchical split in Res2Net that can represent short bursts and long drifts in the same compact block.

Figure 4(c) presents the relation between the triplet margin and correction performance. A very small margin produces insufficient class separation, while an excessive margin forces unrelated windows too far apart and slightly harms coordinate regression. The best validation setting occurs at a margin of 0.35, where validation MAE is 0.32 m and high-risk F1 reaches 0.90. Figure 4(d) compares training stability across five repeated runs. The proposed model shows a standard deviation of 0.018 m in MAE, smaller than the residual CNN and LSTM baselines. This stability is relevant for site-specific training because mine operators may retrain the correction layer after anchor relocation or roadway extension [19].

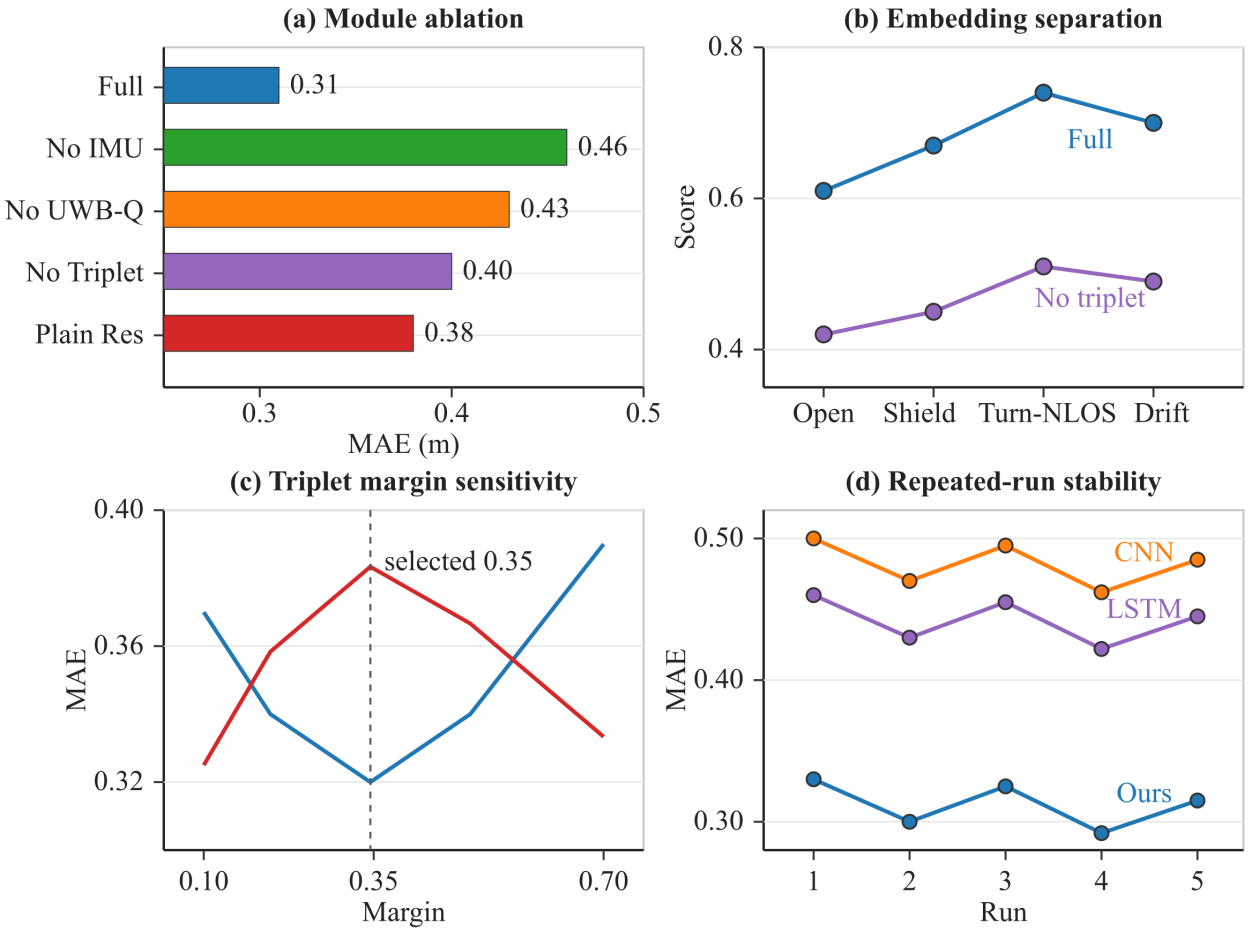


Figure 4. Ablation and training behavior: (a) feature/module ablation; (b) embedding separation score; (c) triplet margin sensitivity; (d) repeated-run stability

Table 2 reports model size, inference latency, and performance under the same hardware. The proposed model has 1.84 million parameters, which is larger than the residual CNN but smaller than the LSTM and transformer-like sequence baseline used in pilot testing. Its 18.7 ms median latency and 24.9 ms 95th-percentile latency leave adequate time for database write-back and visualization.

Table 2. Model complexity, latency, and positioning performance

Model	Parameters (*M*)	Median latency (*ms*)	95th latency (*ms*)	MAE (*m*)	F1
Residual CNN	0.92	12.6	18.2	0.48	0.78
LSTM regressor	2.37	31.5	45.8	0.44	0.82
ResNet correction	1.51	16.8	22.7	0.39	0.85
Res2Net-Triplet	1.84	18.7	24.9	0.31	0.91

Figure 5(a) shows a trajectory segment near a roadway intersection where raw multilateration drifts toward the sidewall during partial shielding; the corrected trajectory remains close to the reference centerline. Figure 5(b) shows the corresponding coordinate-offset vectors, with larger corrections concentrated at the entrance of the occluded segment. Such behavior is desirable because correction effort is applied where measurement inconsistency is observed rather than spread uniformly along the route.

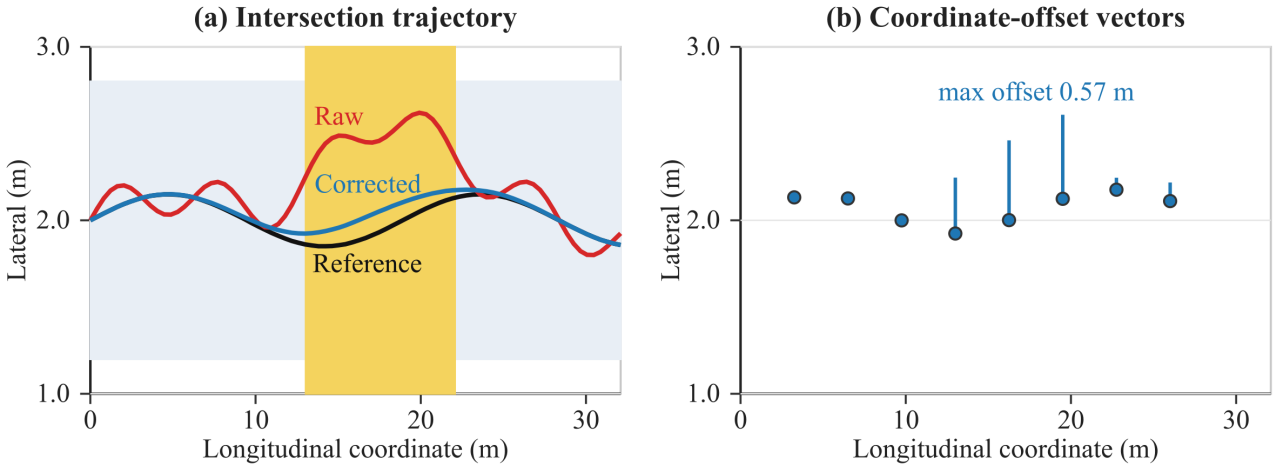


Figure 5. Local trajectory correction: (a) intersection trajectory before and after correction; (b) coordinate-offset vectors along the occluded segment

A trajectory example was selected from a test route not used in training. At the entrance of the occluded region, two anchors are still visible, and the closest anchor has been temporarily blocked by the worker's body. The raw solution is shifted horizontally by the biased range, bringing the multilateration result closer to the roadway boundary. The learned offset gradually increases from 0.18 m to 0.57 m and then decreases after anchor visibility is recovered. This pattern is in line with the input evidence and not imposed by a smoothing rule. A purely smooth trajectory may still be incorrect if it follows the biased sidewall estimate; therefore, the correction model uses channel and inertial context to determine where to focus the offset.

The model is also parameter-efficient. The proposed model is not the smallest network, but it is a relatively compact one that falls between shallow convolution and recurrent sequence models. The residual CNN has a small latency but lacks sufficient representation power for severe NLOS windows. The LSTM baseline has more parameters and a longer latency tail because of the non-parallel nature of recurrent state updates. Res2Net-Triplet adds only a moderate parameter increase over ResNet while improving both tail error and high-risk F1. This trade-off is suitable for monitoring servers that need to update many tags at regular intervals without delaying alarms or database synchronisation.

According to the loss-weight sensitivity test, metric learning should be combined with coordinate regression cautiously. When lambda is set to 0, the model becomes a standard Res2Net regressor and achieves a validation MAE of 0.36 m. When lambda increases to 0.15, validation MAE decreases to 0.32 m and high-risk F1 rises from 0.86 to 0.90. When lambda > 0.35, the embedding branch starts to dominate the shared backbone, and the model separates disturbance categories more strongly than required by the regression target. Therefore, the best configuration treats triplet supervision as a geometric regularizer rather than the primary training goal. This behaviour is in line with the correction task, and the final coordinate offset is the main engineering result.

Ablation studies were carried out with a smaller training set to determine whether the architecture required a large-scale site survey. With only 30% of the training windows, the proposed model has achieved a test MAE of 0.43m; ResNet is at 0.52m and LSTM is at 0.57m. With 60% of the training windows, the proposed model reaches 0.35 m and the performance gap is reduced. Therefore, triplet sampling can be used to address the problem of limited data by increasing the number of informative pairwise relations in a mini-batch. Additional field trajectories will be needed for severe NLOS and equipment-adjacent sections where the disturbance pattern is less frequent.

The magnitude of the error differs by axis. The lateral axis is most likely to be affected by correction because sidewall reflection and body shielding generally bias the range perpendicular to the roadway centreline. The extension is relatively small because the rough UWB geometry has already restricted movement in the straight corridor. The vertical axis has a small absolute error but is still suitable for ramps and uneven floor sections. This axis-level behaviour supports the decision to regress a three-dimensional offset rather than a scalar confidence value. A scalar correction can identify a bad window, but it cannot specify where to move the coordinate back to a reasonable position in the system.

5.3. Robustness, Error Sources, and Operational Analysis

Robustness was evaluated under anchor dropout, altered NLOS severity, and sensor-noise perturbation. Figure 6(a) shows that when one anchor is randomly unavailable, the proposed model increases mean error from 0.31 m to 0.37 m, while ResNet increases from 0.39 m to 0.54 m. With two missing anchors, the proposed model still keeps mean error at 0.52 m because the input mask allows it to interpret missing measurements explicitly. Figure 6(b) compares performance under low, moderate, and severe NLOS levels. Severe NLOS raises the proposed model error to 0.58 m, but the 95th-percentile error remains below 1.26 m, which is lower than the moderate-NLOS result of EKF. This outcome agrees with studies showing that NLOS bias must be modeled as structured disturbance rather than treated as independent noise [20].

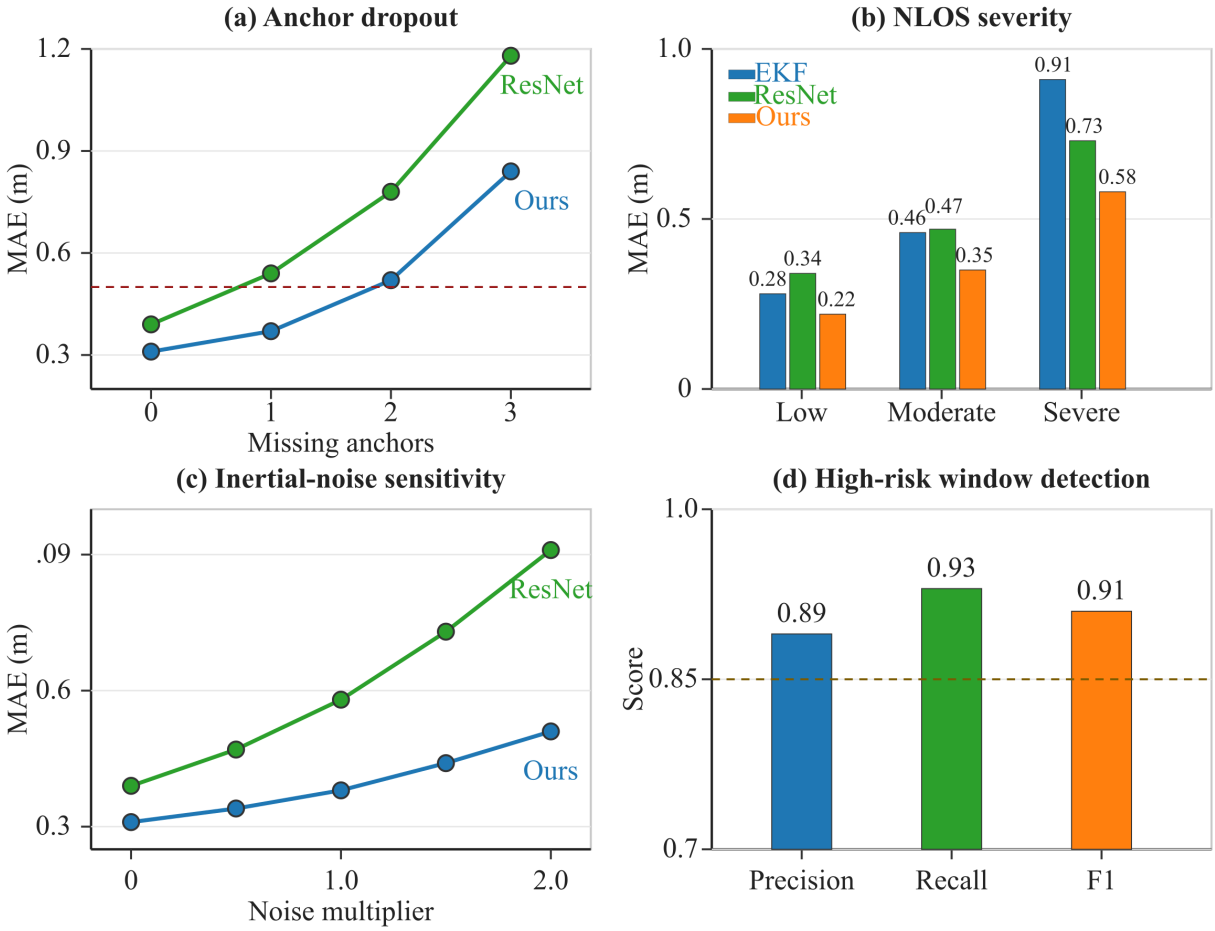


Figure 6. Robustness analysis: (a) anchor-dropout accuracy; (b) NLOS severity response; (c) inertial-noise sensitivity; (d) high-risk window detection

Figure 6(c) reports inertial-noise sensitivity by adding controlled accelerometer and gyroscope perturbations. The proposed model degrades gradually because the UWB quality and mask channels still carry reliable evidence when the inertial stream becomes noisy. Figure 6(d) shows high-risk window detection, where the proposed model achieves 0.91 F1, 0.89 precision, and 0.93 recall. The detection result is derived from corrected-error thresholds and is useful for

raising confidence flags in the monitoring interface. Safety-oriented localization systems should expose uncertainty or risk markers instead of presenting all coordinates as equally reliable [21]. The proposed model supports this practice because its embedding distance and residual magnitude can be converted into a confidence indicator.

Figure 7(a) displays a heat map of correction magnitude over roadway segments and anchor visibility. The strongest correction occurs in junction segments with one or fewer reliable anchors, where median offset reaches 0.71 m, while straight segments with four visible anchors require only 0.22 m correction. Curved segments with two reliable anchors show 0.49 m median offset, which is higher than the 0.41 m value in straight segments under the same anchor count. This pattern indicates that roadway context and measurement availability interact: the same number of visible anchors is less harmful in a straight gallery than at a junction, where small angular errors can redirect a trajectory to a different branch. The heat map also shows that equipment-adjacent areas remain difficult even with three visible anchors, suggesting that reflective objects distort range reliability beyond what anchor count alone can describe.

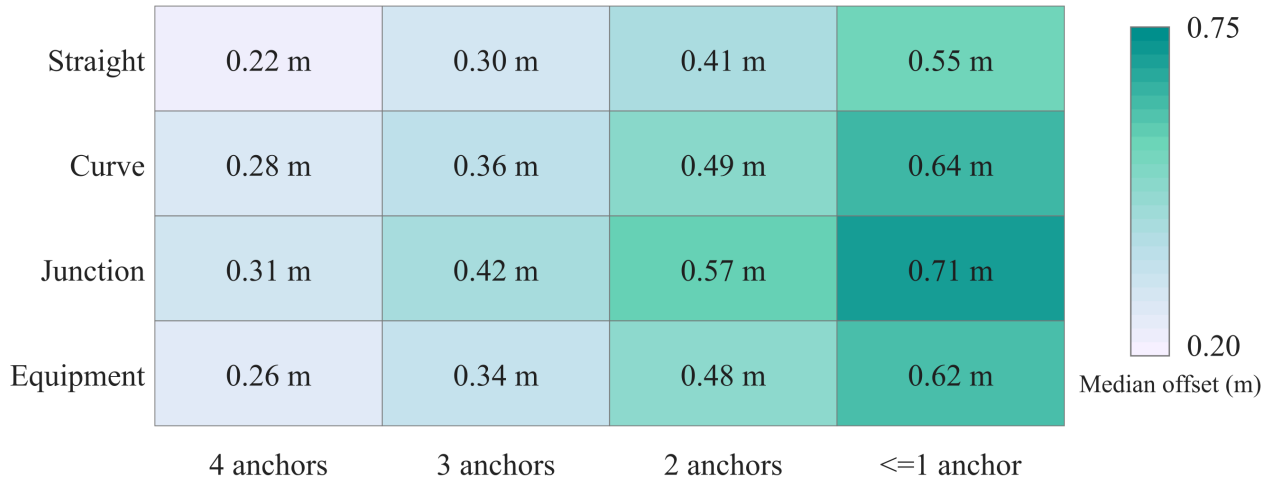


Figure 7. Roadway-context correction magnitude: (a) median offset heat map by segment type and visible-anchor count

Operational performance was analyzed through latency, memory use, update continuity, and high-error sources. Figure 8(a) presents latency distributions for the compared learning models; the proposed model has a median latency of 18.7 ms and remains below the 25 ms processing budget for most windows, while LSTM reaches 31.5 ms and has a longer tail. Figure 8(b) reports model footprint on an independent parameter-memory coordinate system, where the proposed model uses 1.84 million parameters and 33 MB of memory. Figure 8(c) shows continuity-gate rejection rate as a separate percentage chart, rising from 1.8% in open-view windows to 8.9% under severe NLOS. Figure 8(d) presents high-error source composition, where fewer-than-three-anchor windows account for the largest share at 37%, followed by rapid turning at 29%. Real-time deployment work has shown that latency and confidence gating can be as important as nominal positioning error in safety systems [22].

Walking speed, embedding behaviour and layout change were considered secondary robustness factors outside the plotted deployment panels. The typical 0.6-1.4 m/s walking range had a mean error of 0.32-0.35 m, and extremely slow movement increased the error to 0.51 m because inertial periodicity was no longer effective. Open-view walking windows formed the closest embedding group, and while partial shielding and turning NLOS remained adjacent but separable, a similar representation separation has been identified as beneficial for sensor-fusion tasks with heterogeneous disturbance sources [23]. When the test route used anchors shifted by 1.5 m from their training positions, the proposed model increased the mean error by 0.06 m after updating coordinate normalisation, but the increase reached 0.14 m without updating normalisation.

The remaining errors and continuity behaviour were analysed together. A manual review of the 260 high-error windows found that 37% had fewer than three usable anchors, 29% occurred during rapid turning with body shielding, 21% appeared near moving equipment, and the rest were related to reference-label uncertainty or synchronization imperfections. Prior work on industrial positioning also reports that rare geometric degeneracy can dominate the upper tail even when average error appears acceptable [24]. Previous synchronisation studies have also shown that, for accurate correction, only minimal anchor observability and timestamp alignment are needed [25]. Without a continuity gate, 0.7% of the windows produced a correction vector larger than 1.2 m after abrupt anchor recovery; in a server scenario with 80 active tags at 5 Hz, the median inference load is less than two CPU-core equivalents when batching is enabled. Field retraining should be arranged after anchor relocation, roadway extension or significant changes in equipment layout, as the learned correction field is partially site-specific [26].

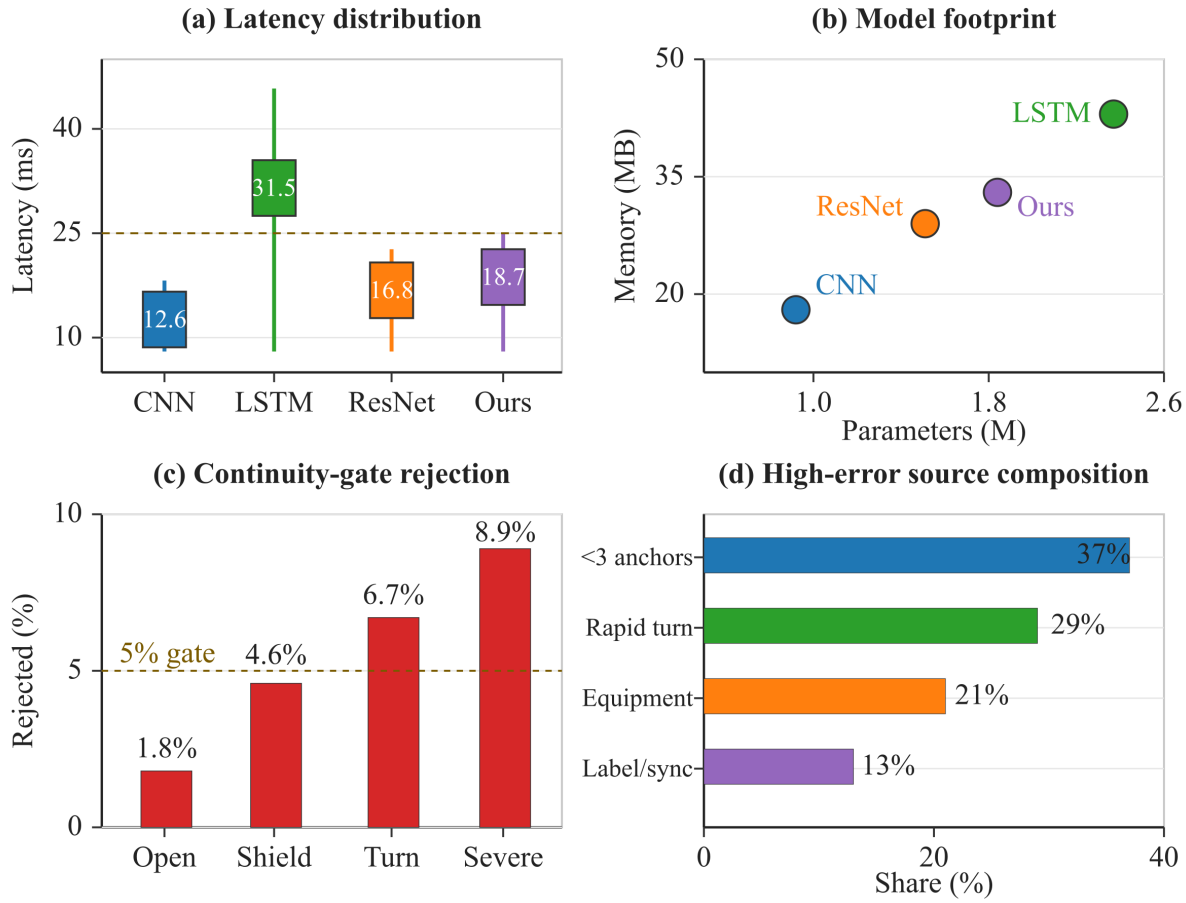


Figure 8. Deployment-oriented analysis: (a) latency distribution; (b) model footprint; (c) continuity-gate rejection rate; (d) high-error source composition

6. Conclusion

This paper proposes a UWB- and inertial-measurement-based Res2Net-Triplet model for underground personnel positioning correction. The method keeps the physical localisation pipeline and learns a coordinate-offset correction from synchronous ranging, channel-quality, inertial, mask and coarse-trajectory windows. Res2Net blocks learn short- and context-dependent disturbance patterns, and triplet supervision orders the embedding space according to correction behaviour. The resulting framework is suitable for roadway environments with NLOS bias, body shielding and inertial drift simultaneously.

The study has some limitations. The experiments use controlled road trajectories and labelled correction windows, so the model's performance will be affected by the quality and diversity of site-specific calibration data. Extremely long UWB outages, abrupt posture changes and severe anchor-geometry degeneration are still difficult cases. The current implementation uses a compact continuity gate instead of a full probabilistic safety layer, and thus has reduced coverage for uncertainty in previously unobserved operating conditions.

Future work will extend the framework to continuous calibration, adaptive anchor-health monitoring and uncertainty-aware deployment. Longer field trials across multiple mine layouts, equipment states and anchor relocation scenarios will be required to evaluate generalization under operational conditions. Future studies should also combine the correction layer with roadway topology constraints and uncertainty calibration so that the monitoring platform can present both corrected coordinates and confidence levels during safety-critical operation.

7. Declarations

7.1 Author Contributions

Kovács István contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Nagy Péter contributes to validation, analysis, investigation, data collection, draft preparation. Farkas Tamás contributes to conceptualization, validation, analysis. All authors have read and agreed with the manuscript before its submission and publication.

7.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

7.3 Institutional Review Board Statement

Not applicable.

7.4 Informed Consent Statement

Not applicable.

7.5 Declaration of Competing Interest

Not applicable.

7.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

8. References

[1] Guo, X., Ansari, N., Hu, F., Shao, Y., Elikplim, N. R., & Li, L. (2019). A survey on fusion-based indoor positioning. IEEE Communications Surveys & Tutorials, 22(1), 566-594. https://doi.org/10.1109/COMST.2019.2951036

[2] Nessa, A., Adhikari, B., Hussain, F., & Fernando, X. N. (2020). A survey of machine learning for indoor positioning. IEEE Access, 8, 214945-214965. https://doi.org/10.1109/ACCESS.2020.3039271

[3] Gao, S. H., Cheng, M. M., Zhao, K., Zhang, X. Y., Yang, M. H., & Torr, P. (2019). Res2net: A new multi-scale backbone architecture. IEEE transactions on pattern analysis and machine intelligence, 43(2), 652-662. https://doi.org/10.1109/TPAMI.2019.2938758

[4] Sidorenko, J., Schatz, V., Bulatov, D., Scherer-Negenborn, N., Arens, M., Hugentobler, U., & Schütz, A. (2020). Error corrections for ultrawideband ranging. IEEE Transactions on Instrumentation and Measurement, 69(11), 9037-9047. https://doi.org/10.1109/TIM.2020.2999244

[5] Elsanhoury, M., Mäkelä, P., Koljonen, J., Välisuo, P., Shamsuzzoha, A., Mantere, T., Elmusrati, M., & Kuusniemi, H. (2022). Precision positioning for smart logistics using ultra-wideband technology-based indoor navigation: A review. IEEE Access, 10, 44413-44445. https://doi.org/10.1109/ACCESS.2022.3169267

[6] Leitch, S. G., Lohan, E. S., Latva-aho, M., & Valkama, M. (2023). On indoor localization using WiFi, BLE, UWB, and IMU technologies. Sensors, 23(23), 9638. https://doi.org/10.3390/s23239638

[7] Li, M., Zhu, H., You, S., & Tang, C. (2020). UWB-based localization system aided with inertial sensor for underground coal mine applications. IEEE Sensors Journal, 20(12), 6652-6669. https://doi.org/10.1109/JSEN.2020.2976097

[8] Capraro, F., Segura, M., & Sisterna, C. (2020). Human real time localization system in underground mines using UWB. IEEE Latin America Transactions, 18(2), 392-399. https://doi.org/10.1109/TLA.2020.9085295

[9] You, W., Li, F., Liao, L., & Huang, M. (2020). Data fusion of UWB and IMU based on unscented Kalman filter for indoor localization of quadrotor UAV. IEEE Access, 8, 64971-64981. https://doi.org/10.1109/ACCESS.2020.2985053

[10] Zhang, Y., Tan, X., & Zhao, C. (2020). UWB/INS integrated pedestrian positioning for robust indoor environments. IEEE Sensors Journal, 20(23), 14401-14409. https://doi.org/10.1109/JSEN.2020.3008815

[11] Ali, R., Liu, R., Nayyar, A., Qureshi, B., & Cao, Z. (2021). Tightly coupling fusion of UWB ranging and IMU pedestrian dead reckoning for indoor localization. IEEE Access, 9, 164206-164222. https://doi.org/10.1109/ACCESS.2021.3132645

[12] Yang, X., Wang, J., Song, D., Feng, B., & Ye, H. (2021). A novel NLOS error compensation method based IMU for UWB indoor positioning system. IEEE Sensors Journal, 21(9), 11203-11212. https://doi.org/10.1109/JSEN.2021.3057369

[13] Feng, D., Peng, J., Zhuang, Y., Guo, C., Zhang, T., Chu, Y., Zhou, X., & Xia, X. G. (2023). An adaptive IMU/UWB fusion

method for NLOS indoor positioning and navigation. IEEE Internet of Things Journal, 10(13), 11414-11428. https://doi.org/10.1109/JIOT.2023.3245144

[14] Liu, C., Kadja, T., & Chodavarapu, V. P. (2022). Experimental evaluation of sensor fusion of low-cost UWB and IMU for localization under indoor dynamic testing conditions. Sensors, 22(21), 8156. https://doi.org/10.1109/JSEN.2022.3182416

[15] Zhang, H., Wang, Q., Li, Z., Mi, J., & Zhang, K. (2023). Research on high precision positioning method for pedestrians in indoor complex environments based on UWB/IMU. Remote Sensing, 15(14), 3555. https://doi.org/10.3390/rs15143555

[16] Hölzke, F., Borstell, H., Golatowski, F., & Haubelt, C. (2023). Pedestrian localization with stride-wise error estimation and compensation by fusion of UWB and IMU data. Sensors, 23(10), 4744. https://doi.org/10.3390/s23104744

[17] De Cock, C., Tanghe, E., Joseph, W., & Plets, D. (2023). Robust IMU-based mitigation of human body shadowing in UWB indoor positioning. Sensors, 23(19), 8289. https://doi.org/10.3390/s23198289

[18] Otim, T., Bahillo, A., Díez, L. E., Lopez-Iturri, P., & Falcone, F. (2020). Towards sub-meter level UWB indoor localization using body wearable sensors. IEEE Access, 8, 178886-178899. https://doi.org/10.1109/ACCESS.2020.3027669

[19] Yang, B., Yang, E., Yu, L., & Niu, C. (2023). Adaptive extended Kalman filter-based fusion approach for high-precision UAV positioning in extremely confined environments. IEEE/ASME Transactions on Mechatronics, 28(1), 543-554. https://doi.org/10.1109/TMECH.2022.3203875

[20] Wang, F., Tang, H., & Chen, J. (2023). Survey on NLOS identification and error mitigation for UWB indoor positioning. Electronics, 12(7), 1678. https://doi.org/10.3390/electronics12071678

[21] Jiang, C., Shen, J., Chen, S., Chen, Y., Liu, D., & Bo, Y. (2020). UWB NLOS/LOS classification using deep learning method. IEEE Communications Letters, 24(10), 2226-2230. https://doi.org/10.1109/LCOMM.2020.2999904

[22] Cui, Z., Gao, Y., Hu, J., Tian, S., & Cheng, J. (2021). LOS/NLOS identification for indoor UWB positioning based on Morlet wavelet transform and convolutional neural networks. IEEE Communications Letters, 25(3), 879-882. https://doi.org/10.1109/LCOMM.2020.3039251

[23] Liu, Q., Yin, Z., Zhao, Y., Wu, Z., & Wu, M. (2022). UWB LOS/NLOS identification in multiple indoor environments using deep learning methods. Physical Communication, 52, 101695. https://doi.org/10.1016/j.phycom.2022.101695

[24] Wei, J., Wang, H., Su, S., Tang, Y., Guo, X., & Sun, X. (2022). NLOS identification using parallel deep learning model and time-frequency information in UWB-based positioning system. Measurement, 195, 111191. https://doi.org/10.1016/j.measurement.2022.111191

[25] Si, M., Wang, Y., Xu, S., Sun, M., & Cao, H. (2023). A lightweight CIR-based CNN with MLP for NLOS/LOS identification in a UWB positioning system. IEEE Communications Letters, 27(5), 1332-1336. https://doi.org/10.1109/LCOMM.2023.3254984

[26] Deng, B., Xu, T., & Yan, M. (2023). UWB NLOS identification and mitigation based on gramian angular field and parallel deep learning model. IEEE Sensors Journal, 23(22), 28513-28525. https://doi.org/10.1109/JSEN.2023.3323564