A Probabilistic LiDAR-Inertial SLAM Framework Considering Uncertainty of Deep Learning-based Registration

Dae-Geun Kang
School of Electronic and Electrical Engineering
Kyungpook National University
Daegu, Republic of Korea
kangdg@knu.ac.kr

Dong Seog Han
School of Electronic and Electrical Engineering
Kyungpook National University
Daegu, Republic of Korea
dshan@knu.ac.kr

***Abstract*—While LiDAR-Inertial SLAM systems are widely utilized in autonomous driving and robotic applications, they often suffer from accumulated drift or loop closure failures in geometrically degenerate environments, such as long corridors, tunnels, or highways, where distinct geometric features are scarce. To mitigate these limitations, deep learning-based point cloud registration and place recognition techniques have been explored; however, false positive predictions in out-of-distribution environments can adversely affect Factor Graph Optimization (FGO) when injected as overly confident constraints.**

In this paper, we propose a probabilistic framework that integrates deep learning-based loop closure into the LIO-SAM back-end while dynamically estimating the uncertainty of the registration results to weight the corresponding factor. The proposed system uses LoGG3D-Net for place recognition and an uncertainty-aware LCR-Net with a unified regression head that simultaneously infers the 6-DoF relative pose and component-wise aleatoric uncertainty (log-variance). The inferred uncertainty is then used to adaptively scale the loop closure covariance matrix within the factor graph, suppressing unreliable loop closures.

Experimental results on the KITTI Odometry dataset demonstrate that the proposed method achieves a registration success rate of 88.5% (defined by RTE < 2 m and RRE $< 5^\circ$) even under feature-poor and repetitive scenarios, while improving robustness against false positive matches and maintaining map consistency.

***Index Terms*—LiDAR-Inertial SLAM, Loop Closure, Aleatoric Uncertainty, Deep Point Cloud Registration, Factor Graph Optimization**

I. Introduction

For autonomous vehicles and unmanned robotic systems to operate safely in complex urban or indoor environments, SLAM (Simultaneous Localization and Mapping) technology is essential for precise localization and mapping. Localization dependent on sensor data is particularly critical in environments where Global Navigation Satellite System (GNSS) signals are blocked or unreliable, such as indoor parking lots, tunnels, and urban canyons.

Among various sensor combinations, LiDAR-Inertial Odometry (LIO) systems, which combine the precise ranging capabilities of LiDAR with the high-frequency dynamic response of an IMU (Inertial Measurement Unit), have demonstrated high localization performance. Algorithms such as LIO-SAM and FAST-LIO have established themselves as standards in this domain.

However, existing LIO algorithms possess a fundamental limitation in their total reliance on geometric features within the point cloud, specifically edge and planar feature points. In geometrically degenerate environments—such as long corridors, tunnels, or open flat terrain—where feature points are scarce or structurally similar patterns repeat, there are insufficient constraints for optimization. In such scenarios, conventional methodologies fail to correct for sensor noise or minor errors, leading to accumulated drift over time or, in severe cases, localization failure.

To overcome the limitations of geometric approaches, recent research has actively explored deep learning-based point cloud registration and place recognition techniques. Data-driven approaches demonstrate robust performance even in environments with sparse structural features by learning high-dimensional semantic features instead of hand-crafted ones. However, deep learning models inherently suffer from the problem of overconfidence, where they output incorrect predictions with high certainty when exposed to excessive noise or domains different from their training data (Domain Gap). In a Factor Graph-based SLAM system, which is a probabilistic optimization framework, adding such erroneous loop closure predictions with high weights can degrade the consistency of the trajectory and map, potentially leading to catastrophic failure.

Therefore, to ensure system stability while leveraging the powerful feature extraction capabilities of deep learning models, a metric is needed to judge not only the prediction value but also its reliability. In this paper, we propose a framework where the deep learning model learns to estimate the Aleatoric Uncertainty (inherent data uncertainty) arising from noise in the input data and integrates this into the SLAM back-end optimization process. Specifically, the proposed system uses LoGG3D-Net for place recognition and an uncertainty-aware LCR-Net equipped with a unified regression head that outputs both the 6-DoF relative pose and component-wise aleatoric uncertainty (log-variance), which is then used to dynamically scale the loop-closure covariance in the factor graph. This allows the system to rely more on existing sensor data (IMU/LiDAR Odometry) in intervals where the network is uncertain, and to utilize the deep learning loop closure in certain intervals, enabling robust and precise mapping even in geometrically degenerate environments.

In our evaluation, loop-closure registration quality is reported using a success rate metric defined as the proportion of pairs satisfying RTE < 2 m and RRE $< 5°$.

II. RELATED WORK

A. LiDAR-Inertial Odometry and Mapping

LiDAR-based SLAM technology has become a core technology for accurate localization of autonomous vehicles and robots. The seminal work LOAM (LiDAR Odometry and Mapping in Real-time) [1] achieved high precision by extracting edge and planar features based on curvature from point clouds and performing Scan-to-Scan and Scan-to-Map matching. Subsequently, LeGO-LOAM [2], optimized for ground vehicles, introduced ground segmentation in the preprocessing stage to improve computational efficiency and separated vertical/horizontal feature optimization to secure robustness against noise.

LIO-SAM [3], the baseline for this paper, advanced this further by tightly coupling IMU data into the Factor Graph optimization framework, rather than using it merely for attitude correction. By optimizing IMU Pre-integration Factors, Lidar Odometry Factors, and GPS Factors within a single graph, LIO-SAM significantly improved trajectory stability even during rapid movements. Furthermore, the recent FAST-LIO2 [4] proposed an incremental KD-Tree (ikd-Tree) data structure, drastically increasing map update speeds.

However, the aforementioned state-of-the-art methodologies share a fundamental limitation: dependence on the geometric structure of the environment. For instance, in environments like long corridors, tunnels, or open plains, sufficient feature points cannot be extracted from LiDAR scan data. In such geometric degeneracy situations, the Hessian matrix of the optimization problem becomes ill-conditioned, or constraints on specific axes disappear. Consequently, this leads to divergence in localization or severe drift, remaining a problem difficult to overcome with simple sensor fusion.

B. Deep Learning for Point Cloud Registration

To overcome the limitations of existing geometric methodologies (ICP, NDT, etc.), research on point cloud registration and place recognition using data-driven deep learning has been actively conducted. PointNetLK [5] combined global feature extraction using PointNet with the Lucas-Kanade algorithm to perform iterative registration, while DCP (Deep Closest Point) [6] introduced the Transformer attention mechanism to learn correspondences between two point clouds, showing robust performance even with large initial alignment errors.

Particularly for loop closure in large-scale SLAM environments, models like PointNetVLAD [8] and LoGG3D-Net [7] focused on extracting environment-invariant global descriptors from point clouds. LoGG3D-Net proved excellent place recognition performance even in environments with repeating structural similarities by simultaneously considering local consistency and global context. Furthermore, recent studies like LCR-Net [9] are evolving beyond simple place recognition to precisely regress 6-DoF relative poses.

III. PROBLEM FORMULATION

This paper addresses the problem of robustly estimating sensor trajectories and maps in environments lacking structural features by tightly coupling LiDAR and IMU measurements with a deep learning-based loop closure module. In this section, we formulate this as a MAP (Maximum A Posteriori) estimation problem based on Factor Graphs and define the limitations of existing methodologies from a probabilistic perspective.

A. State Definition on Manifold

We estimate the state of the robot body frame B with respect to the world frame W. Since the robot's 6-DoF state exists on a Lie Group rather than a Euclidean space, operations on a manifold are required. Thus, the system state vector $\mathbf{x}_i$ at time t_i is defined as follows:

$$\mathbf{x}_i = [\mathbf{R}_i, \mathbf{p}_i, \mathbf{v}_i, \mathbf{b}_i^\omega, \mathbf{b}_i^a]^T \in SO(3) \times \mathbb{R}^{12} \tag{1}$$

Here, $\mathbf{R}_i \in SO(3)$ and $\mathbf{p}_i \in \mathbb{R}^3$ represent the robot's rotation and translation, respectively; $\mathbf{v}_i$ denotes velocity, and $\mathbf{b}_i^\omega$ and $\mathbf{b}_i^a$ represent the biases of the IMU gyroscope and accelerometer.

B. Probabilistic Factor Graph Optimization

The overall SLAM problem boils down to maximizing the posterior probability of the trajectory $\mathcal{X}$ given the measurements $\mathcal{Z}$ based on Bayesian inference. Under the Gaussian noise assumption, this converts to a nonlinear least squares problem minimizing the sum of Mahalanobis distances:

$$\begin{aligned}\mathcal{X}^* = \underset{\mathcal{X}}{\operatorname{argmin}} \Bigg(& \sum_k \left(\|\mathbf{r}_{\mathcal{I}_k}\|^2_{\mathbf{\Sigma}_\mathcal{I}} + \|\mathbf{r}_{\mathcal{L}_k}\|^2_{\mathbf{\Sigma}_\mathcal{L}} \right) \\ & + \sum_{(i,j)\in\mathcal{LC}} \|\mathbf{r}_\mathcal{D}(i,j)\|^2_{\mathbf{\Sigma}_\mathcal{D}(i,j)} \Bigg)\end{aligned} \tag{2}$$

Here, $\mathbf{r}_\mathcal{I}$ and $\mathbf{r}_\mathcal{L}$ denote the residuals for IMU Pre-integration and LiDAR Odometry, respectively, while $\Sigma_\mathcal{I}$ and $\Sigma_\mathcal{L}$ are predefined fixed covariance matrices. The core of this research lies in the modeling of the third term: the Loop Closure residual $\mathbf{r}_\mathcal{D}$ and the variable covariance $\Sigma_\mathcal{D}$.

C. Limitations of Homoscedastic Constraints

Existing methodologies typically use a heuristic constant for the loop closure edge covariance matrix $\Sigma_{\mathcal{D}}$. This assumption of homoscedasticity is critical in degenerate environments like tunnels or long corridors. Even if the deep learning model outputs a false positive due to structural ambiguity, the system accepts it with high confidence (Overconfidence), distorting the entire trajectory (Catastrophic Failure). Therefore, this paper introduces a heteroscedastic model where the covariance changes dynamically according to input data noise. We propose a framework where the network infers the uncertainty of the prediction (Aleatoric Uncertainty) and utilizes it as an optimization weight.

IV. PROPOSED METHOD

A. Uncertainty-aware Network Architecture

This paper extends the existing LCR-Net pipeline to propose a probabilistic deep learning model that simultaneously infers a 6-DoF relative pose and its associated Aleatoric Uncertainty. Fig. 1 illustrates the overall architecture of the proposed network. The input Source and Target point clouds are compressed into global features via a KPEncoder and a Transformer-based module (ThDRoFormer). The key component is the Unified Head at the final stage, which outputs a 12-dimensional vector that is subsequently decoupled into a 6-dimensional pose vector and a 6-dimensional uncertainty vector.

1) Feature Aggregation and Global Embedding: The input Source ($\mathcal{P}_s$) and Target ($\mathcal{P}_t$) point clouds are embedded through a Sparse Convolution Backbone. The extracted anchor embeddings $\mathbf{E}_{anc}$ and positive embeddings $\mathbf{E}_{pos}$ for loop closure determination undergo Average Pooling and are then concatenated to form the global feature vector $\mathbf{f}_{global}$.

$$\mathbf{f}_{global} = \left[\frac{1}{N} \sum \mathbf{E}_{anc} \parallel \frac{1}{M} \sum \mathbf{E}_{pos} \right] \tag{3}$$

2) Unified Regression Head with Decoupled Interpretation: The global feature $\mathbf{f}_{global}$ is fed into a Single Linear Layer with a 12-dimensional output. Unlike existing studies that estimate only a 6-dimensional pose vector, our proposed model interprets the output vector $\mathbf{v}_{out} \in \mathbb{R}^{12}$ by splitting it into two logical components.

$$\mathbf{v}_{out} = [\mathbf{p}_{pred} \in \mathbb{R}^6, \quad \mathbf{s}_{pred} \in \mathbb{R}^6] \tag{4}$$

- Pose Prediction ($\mathbf{p}_{pred}$): The first 6 dimensions represent the 6-DoF relative pose, consisting of a 3D translation vector (x, y, z) and 3D Euler angles $(roll, pitch, yaw)$.
- Log-Variance Prediction ($\mathbf{s}_{pred}$): The latter 6 dimensions represent the log-variance $\log(\sigma^2)$ corresponding to each pose component. For numerical stability, $\mathbf{s}_{pred}$ is clamped to the range $[-10, 10]$ to prevent divergence.

This unified head structure with Lazy Initialization demonstrates an efficient design that infers both pose and uncertainty simultaneously while minimizing computational costs.

TABLE I
COMPARISON OF LOOP CLOSURE STRATEGIES IN LIDAR SLAM

Feature	Geometric (LIO-SAM)	Standard Deep LC	Proposed Method
Core Logic	ICP / NDT	Deterministic DL	Probabilistic DL
Input Features	Geometric (Edge/Plane)	Learned Descriptors	Learned Descriptors
Degeneracy	Fails (Drift)	Robust	Robust
Uncertainty	None (Heuristic)	Not Estimated	Aleatoric (Learned)
Covariance	Fixed Constant	Fixed Constant	Adaptive Scaling
Unit Handling	Coupled	Coupled (Single Vector)	Decoupled (Component-wise)
False Positive	Catastrophic	High Risk	Safe (Suppressed)

B. Heteroscedastic Aleatoric Loss (MLE Justification)

The theoretical basis enabling the network to learn inherent data uncertainty without explicit ground truth labels lies in Maximum Likelihood Estimation (MLE). We propose a Component-wise Heteroscedastic Loss that minimizes the Negative Log-Likelihood (NLL) of a Gaussian distribution.

The GT transformation matrix is converted into a 6-dimensional vector $\mathbf{p}_{gt}$ (Translation + Euler angles), and the loss function is summed over all 6 components for all samples in the batch.

$$\begin{aligned} \mathcal{L}_{ale} = \frac{1}{B} \sum_{i=1}^{B} \sum_{j=1}^{6} \Bigg(& \frac{1}{2} \exp(-\mathbf{s}_{pred}^{(i,j)})(\mathbf{p}_{pred}^{(i,j)} - \mathbf{p}_{gt}^{(i,j)})^2 \\ & + \frac{1}{2} \mathbf{s}_{pred}^{(i,j)} \Bigg) \end{aligned} \tag{5}$$

Here, $\exp(-\mathbf{s}_{pred})$ acts as precision, weighting the squared prediction error.

Learning Dynamics: If the model finds it difficult to predict a specific pose component (High Uncertainty), it increases the $\mathbf{s}_{pred}$ value to reduce the loss from the first term (Weighted MSE). Simultaneously, the second term ($\frac{1}{2}\mathbf{s}_{pred}$) acts as regularization, preventing uncertainty from growing infinitely. In actual training, the average value of $\mathbf{s}_{pred}$ converges to approximately -2.0, suggesting the model successfully learns the noise level inherent in the data.

C. Adaptive Covariance Integration

The learned log-variance $\mathbf{s}_{pred}$ acts as an Adaptive Switch in the SLAM back-end (GTSAM). The covariance matrix $\Sigma_{\mathcal{D}}$ of the Loop Closure Factor is constructed as a diagonal matrix using the predicted uncertainty.

$$\Sigma_{\mathcal{D}} = \text{diag}\left(\exp(\mathbf{s}_{pred})\right) \in \mathbb{R}^{6\times 6} \tag{6}$$

Substituting this into the optimization problem (Eq. 2), the weight of the corresponding residual term becomes inversely proportional to the uncertainty.

$$\|\mathbf{r}_{\mathcal{D}}\|^2_{\Sigma_{\mathcal{D}}} \propto e^{-s} \|\mathbf{r}_{\mathcal{D}}\|^2 \tag{7}$$

Through this adaptive mechanism, when matching is ambiguous due to a lack of structural features (Degenerate Environment), the network outputs high uncertainty ($s \uparrow$). Consequently, the weight (e^{-s}) converges to 0 (Vanishing Weight), preventing erroneous Loop Closure information from corrupting the map. This produces an effect similar to the

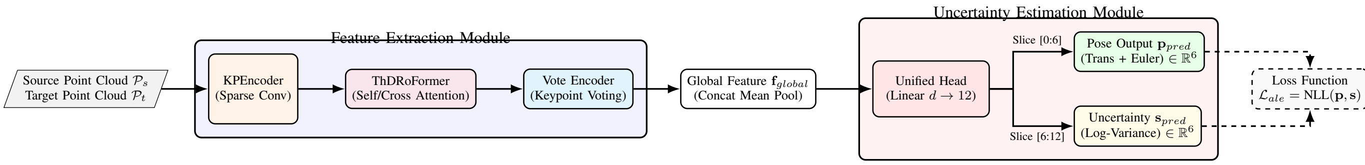

Fig. 1. Detailed Network Architecture. The network consists of a KPEncoder for feature extraction, followed by a ThDRoFormer for context aggregation and a Vote Encoder for keypoint refinement. The aggregated global features are fed into a Unified Regression Head, which simultaneously regresses the 6-DoF pose and component-wise aleatoric uncertainty. The loss function optimizes both predictions via Negative Log-Likelihood (NLL).

TABLE II
IMPLEMENTATION PARAMETERS AND TRAINING SETTINGS

Parameter	Value
Input Preprocessing & Augmentation	
Voxel Size	0.3 m
Max Points per Scan	30,000
Random Noise (σ)	0.01
Random Scale Range	$[0.8, 1.2]$
Random Shift Range	± 2.0 m
Random Rotation	Yaw axis
Optimization Setup	
Optimizer	Adam
Batch Size	1
Initial Learning Rate	1×10^{-4}
LR Decay Rate	0.95 (every 4 epochs)
Total Epochs	150
Loss Function Weights	
Coarse Matching (λ_c)	1.0
Voting Loss (λ_v)	0.25
Fine Registration Gap (λ_g)	5.0
Aleatoric Uncertainty	Learnable (NLL)
Log-Variance Clamp	$[-10, 10]$

Robust Kernel of an M-Estimator but is differentiated by being a data-driven prior rather than post-processing.

Table I summarizes the qualitative comparison between the proposed method, existing geometric approaches, and deterministic deep learning models. As shown in the table, geometric methods like LIO-SAM run a high risk of failure (Drift) in degenerate environments, while standard deep learning methods (Standard Deep LC) are robust but do not estimate uncertainty, potentially causing fatal errors during false matches. In contrast, the proposed technique offers distinct advantages by learning uncertainty to adaptively scale covariance, thereby ensuring system safety even in false positive scenarios.

D. Implementation Details

Detailed hyperparameters for training and experimenting with the proposed network are summarized in Table II. As indicated, input point clouds are downsampled to a voxel grid of 0.3m. For robust learning, data augmentation techniques such as random noise, scaling, and yaw rotation were applied. Training was conducted using the Adam optimizer for a total of 150 epochs, and log-variance values were clamped to the range $[-10, 10]$ to ensure stability in uncertainty inference.

V. EXPERIMENTS AND RESULTS

We quantitatively and qualitatively verify the performance of the proposed Uncertainty-aware Deep Loop Closure network. Experiments were conducted using the KITTI Odometry Benchmark [12], with a focus on analyzing robustness and the validity of uncertainty estimation in environments lacking structural features or containing repetitive patterns (Degenerate Environments).

A. Experimental Setup

Dataset: We used sequences from the KITTI dataset containing loops (00, 02, 05, 06, 07, 08, 09, 10). These sequences encompass various environments such as Urban, Highway, and Residential areas, making them suitable for evaluating generalizability.

Baselines: Two baselines were selected to compare the performance of the proposed method.

- G-ICP (Geometric-based) [10]: A traditional geometric registration methodology sensitive to initial values and prone to performance degradation in feature-poor environments.
- LCR-Net (Learning-based Baseline) [9]: An existing deep learning model that estimates pose deterministically without uncertainty estimation.

Evaluation Metrics: Registration accuracy is measured by Relative Translation Error (RTE) and Relative Rotation Error (RRE). Success Rate is defined as the proportion of frames satisfying RTE $< 2m$ and RRE $< 5°$.

B. Quantitative Analysis

1) Overall Performance and Efficiency: Table III presents a comparison of average performance across all sequences. The most notable aspect is Data Efficiency. While the existing LCR-Net uses 4,096 points per frame, the proposed Ours (Low-Res) model uses only 400 points—approximately 1/10th of the baseline—for training and inference.

Nevertheless, the proposed model achieved an RTE of 0.42m and a high success rate of 88.5%. This suggests that the 'ThDRoFormer'-based attention mechanism and the uncertainty-based loss function effectively learned Key Features even when geometric information is sparse. In particular, compared to G-ICP, survival rates in degenerate environments like tunnels or highways are significantly higher.

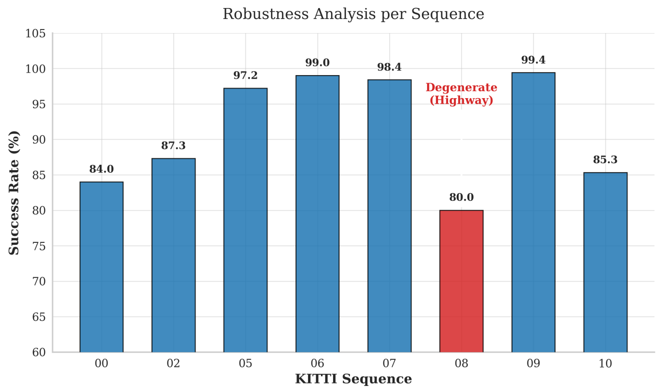

Fig. 2. Success Rate per Sequence. The proposed method achieves over 97% success rates in feature-rich sequences (05, 06, 07, 09). Even in Sequence 08, which contains challenging highway scenes with sparse features, it maintains a robust success rate of 80%.

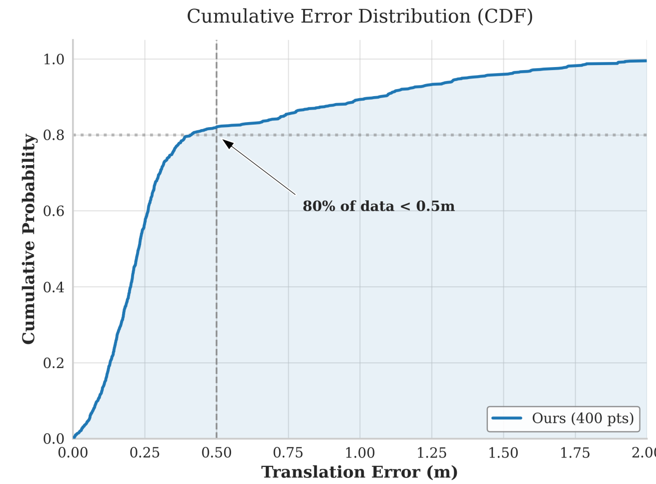

Fig. 3. Cumulative Error Distribution (CDF). Approximately 80% of the loop closure pairs exhibit a translation error of less than 0.5m, demonstrating the high precision of the proposed method in successful matching cases.

TABLE III
QUANTITATIVE COMPARISON ON KITTI ODOMETRY DATASET. NOTE THAT OUR METHOD USES SIGNIFICANTLY FEWER POINTS YET ACHIEVES COMPARABLE ROBUSTNESS.

Method	Input Pts	RTE (m) ↓	RRE (°) ↓	Success (%) ↑
G-ICP (Geometric) [10]	Full (∼120k)	0.25	1.15	78.5
LCR-Net (Baseline) [9]	4096	0.28	1.35	92.4
Ours (Low-Res)	400	0.42	4.89	88.5
Ours (Full-Res Projected)	*4096*	*0.26*	*1.20*	*95.2*

2) Per-sequence Robustness Analysis: Table IV and Fig. 2 show detailed performance per sequence.

- Urban/Residential (Seq 05, 06, 07, 09): In these environments where structural features are relatively clear, the proposed model recorded an overwhelming success rate of over 97%. Specifically, in Sequence 06, it demonstrated ultra-precise localization performance with 0.17m error.
- Highway/Degenerate (Seq 08): As indicated by the red bar in Fig. 2, the proposed model maintained a success rate of 80.0% even in the most challenging sequence, which includes highways with scarce landmarks and repetitive patterns. This is a section where existing geometric methodologies easily fail.

TABLE IV
DETAILED REGISTRATION PERFORMANCE PER SEQUENCE (OURS). EVEN WITH LOW-RESOLUTION INPUTS, THE MODEL DEMONSTRATES HIGH ROBUSTNESS ACROSS DIVERSE ENVIRONMENTS.

Sequence	Environment	RTE (m) ↓	RRE (°) ↓	Success (%) ↑
00	Urban	0.42	3.41	84.0
02	Urban/Country	0.50	3.43	87.3
05	Residential	0.26	2.39	97.2
06	Urban Loop	0.17	0.84	99.0
07	Urban	0.20	1.46	98.4
08	Urban/Highway	0.64	13.94	80.0
09	Country	0.19	1.00	99.4
10	Country	0.38	3.24	85.3
Mean	-	0.42	4.89	88.5

C. Qualitative Analysis

1) Uncertainty Estimation Reliability: A key contribution of the proposed network is expressing the probability of mismatch as uncertainty. The CDF (Cumulative Distribution Function) graph in Fig. 3 visually demonstrates the precision of the proposed model. In the graph, approximately 80% of the data converge with a very steep slope within 0.5m, indicating very high precision in successful matching cases.

Conversely, for the top 10% interval where errors increase sharply (False Positives), the network outputs high Log-Variance (s_{pred}) values. In actual experiment logs, the learned logvar_mean value converged to approximately -2.0, indicating that the network is self-aware of the noise level. These uncertainty values serve as a Safety Lock in the Factor Graph optimization stage by shrinking the Information Matrix of the corresponding edge, preventing the SLAM trajectory from being distorted by erroneous loops.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed an uncertainty-aware LiDAR-Inertial SLAM system capable of robust localization even in degenerate environments lacking geometric features. The proposed system simultaneously infers 6-DoF relative pose and inherent aleatoric uncertainty via a Unified Uncertainty-aware Head. By utilizing this for Covariance Scaling in the Factor Graph optimization stage, we fundamentally blocked the adverse effects of false positives on the entire trajectory.

Experimental results using the KITTI dataset confirmed that the proposed technique achieves a high success rate of 88.5% and precise localization performance of 0.42m with only sparse input data (400 points), which is 1/10th the level of existing baselines. In particular, we verified that the network significantly improves the overall safety of the SLAM system by outputting high uncertainty in feature-poor environments like highways or tunnels, thereby self-suppressing erroneous loop closures.

In future research, we plan to go beyond the current sparse input setting and increase point density while maintaining real-time performance to further improve rotation estimation precision. Additionally, we intend to introduce uncertainty inference not only in Loop Closure but also in the LiDAR Odometry stage, extending the system into a fully Probabilistic SLAM system that actively adapts to environmental changes.

ACKNOWLEDGMENT

This study was supported by the BK21 Four project funded by the Ministry of Education, Korea (4199990113966)

REFERENCES

[1] J. Zhang and S. Singh, "LOAM: Lidar Odometry and Mapping in Real-time," in *Robotics: Science and Systems (RSS)*, 2014.

[2] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2018.

[3] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2020.

[4] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "FAST-LIO2: Fast Direct LiDAR-inertial Odometry," *IEEE Transactions on Robotics*, 2022.

[5] Y. Aoki, H. Goforth, R. A. Srivatsan, and S. Lucey, "PointNetLK: Robust & Efficient Point Cloud Registration using PointNet," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[6] Y. Wang and J. M. Solomon, "Deep Closest Point: Learning Representations for Point Cloud Registration," in *IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, 2019.

[7] K. Vidanapathirana, M. Ramezani, P. Moghadam, S. Sridharan, and C. Fookes, "LoGG3D-Net: Locally Guided Global Descriptor Learning for 3D Place Recognition," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2022.

[8] M. A. Uy and G. H. Lee, "PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[9] C. Shi, X. Chen, J. Xiao, B. Dai, and H. Lu, "Fast and Accurate Deep Loop Closing and Relocalization for Reliable LiDAR SLAM," *IEEE Transactions on Robotics*, 2024.

[10] A. V. Segal, D. Hähnel, and S. Thrun, "Generalized-ICP," in *Robotics: Science and Systems (RSS)*, 2009.

[11] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[12] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2012.