ICRA2026 Late Breaking Results Poster presented at
2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Learning Traversability Cost Maps with Decomposed Uncertainties via Continuous-state MEDIRL

Gwanhyeong Song[1], Dongjae Lee[1], and Ayoung Kim[1*]

Abstract— **Accurate traversability assessment is critical for mobile robot motion planning, yet sensor occlusions and model limitations often compromise cost map reliability. Therefore, analyzing spatial uncertainty is essential for robust risk management. We propose a novel Maximum Entropy Deep Inverse Reinforcement Learning (MEDIRL) framework that learns a traversability cost map while explicitly disentangling aleatoric and epistemic uncertainties. Aleatoric uncertainty is captured via latent sampling in a Conditional Variational Autoencoder, while epistemic uncertainty is estimated using a decoder ensemble. For kinematic fidelity, we introduce efficient continuous-state rollouts utilizing precomputed transition grids and bilinear interpolation. Fusing camera and LiDAR features, our model achieves stable convergence guided by a novel margin loss. Results demonstrate that learned state visitation frequencies match expert trajectories, and the decomposed uncertainties effectively identify high-risk terrains, providing a crucial foundation for safer autonomous navigation.**

I. Introduction & Related Work

Safe and effective autonomous navigation relies heavily on a mobile robot's ability to evaluate the navigability of its surroundings. While Maximum Entropy Deep Inverse Reinforcement Learning (MEDIRL) [1] has proven powerful for deriving cost maps directly from expert demonstrations, bridging the gap between these learned maps and reliable downstream motion planning remains a significant challenge.

First, intertwined prediction errors hinder reliable risk assessment. Sensor occlusions cause data-inherent noise, while out-of-distribution terrains expose a lack of model knowledge. Therefore, explicitly decoupling them into aleatoric and epistemic uncertainties is imperative for context-aware behaviors, such as cautious exploration or strict risk aversion.

Second, conventional reinforcement learning formulations lack kinematic fidelity. They typically rely on discrete grid-based Markov Decision Process (MDP), which fail to capture the continuous, non-holonomic kinematics of real robots. This approximation creates a critical gap between the modeled environment and real-world deployment.

To overcome these limitations, we introduce an advanced MEDIRL architecture designed to explicitly separate aleatoric and epistemic uncertainties. We integrate a Conditional Variational Autoencoder (CVAE) [2] to model aleatoric variance and a decoder ensemble for epistemic uncertainty. To stabilize training given this expanded network

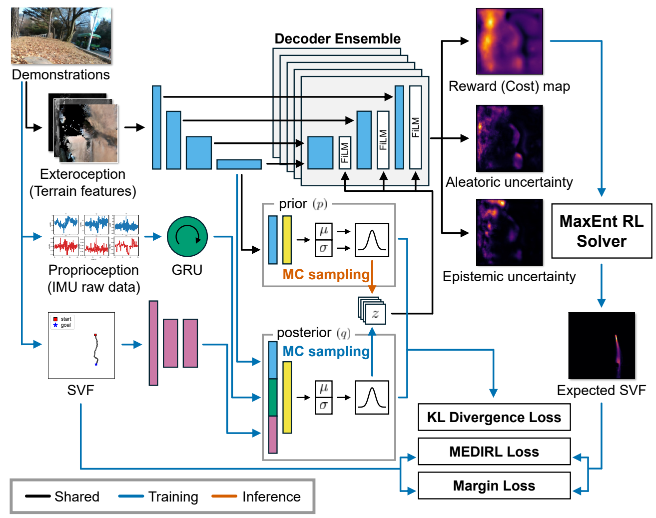

Fig. 1: Overview of the proposed uncertainty-aware MEDIRL framework for disentangling uncertainties in cost map learning.

capacity, we incorporate a margin loss function that penalizes generated trajectories with lower costs than the expert baselines, thereby ensuring robust convergence. For kinematic realism, we introduce continuous-state rollouts using precomputed transition matrices and bilinear interpolation for computational efficiency.

Our main contributions are summarized as follows:

- Extending the MEDIRL framework by incorporating a CVAE and a decoder ensemble to disentangle aleatoric and epistemic uncertainties in traversability cost maps.
- Enhancing kinematic fidelity by formulating the robot state in a continuous (x, y, θ) space.
- Integrating a margin loss to force convergence by penalizing generated rollouts that achieve lower costs than the expert demonstrations.

II. Method

A. System Overview

As illustrated in Fig. 1, our framework employs a teacher-student CVAE to learn a traversability cost map. During training, the posterior (*teacher*) encodes proprioceptive IMU data and LiDAR-IMU SLAM-derived expert State Visitation Frequency (SVF) into a latent distribution, while the prior (*student*) relies solely on exteroceptive features. A KL divergence loss aligns these distributions, enabling cost map generation from solely exteroceptive data during inference.

The sampled latent variable modulates a decoder ensemble via FiLM layers to predict the cost and uncertainty map. A MaxEnt RL solver computes the expected SVF using

†This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2023-00241758).
[1]G. Song, D. Lee and A. Kim are with the Department of Mechanical Engineering, SNU, Seoul, S. Korea [skh8464, pur22, ayoungk]@snu.ac.kr

TABLE I: Calculation methods for multimodal terrain features.

Category	Feature	Calculation
Color	Mean RGB	$\mathbb{E}[\mathbf{p}_c]$ $(c \in \{r, g, b\})$
	Variance RGB	$\mathrm{Var}[\mathbf{p}_c]$ $(c \in \{r, g, b\})$
Geometric	Mean z	$\mathbb{E}[\mathbf{p}_z]$ (below 0.1-quantile)
	Terrain height	Minimum & Gaussian filter
	Slope	$\lVert\nabla \mathbf{p}_z\rVert$
	Obstacle	$\max(\mathbf{p}_z)-$Terrain height
	Roughness	$\rho_{\mathrm{SVD}} = \lambda_3 / \sum_i \lambda_i$
	Point density	Point count per grid

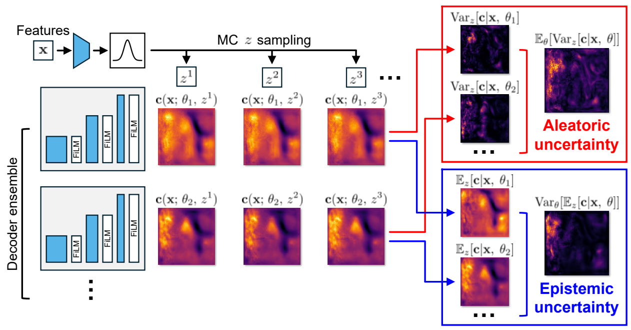

Fig. 2: Illustration of the predictive uncertainty decomposition.

continuous-state rollouts. The architecture is trained end-to-end using the KL divergence, MEDIRL, and margin loss.

B. Terrain Feature Extraction

To generate the input representation for our model, we first colorize the 3D pointcloud from the LiDAR by projecting it onto the camera views. This painted pointcloud is then discretized into 2D grid cells with a resolution of $0.1\,\mathrm{m} \times 0.1\,\mathrm{m}$. For each cell, we extract a 12-channel feature map [3] as summarized in Table. I.

C. MDP Formulation

- **State space**: (x, y, θ) $(x, y \in [-4\,\mathrm{m}, 4\,\mathrm{m}]$, $\theta \in [0, 2\pi])$.
- **Action space**: $\mathcal{A}$ = {straight, soft left, soft right, in-place left, in-place right, stop} $(a = (v_x, \omega_z))$.
- **State transition**: Unicycle kinematic model.

D. Uncertainty Decomposition

As depicted in Fig. 2, aleatoric uncertainty is computed as the variance of predictions derived Monte Carlo sampling in the latent z space, whereas epistemic uncertainty is estimated by the variance across the decoder ensemble.

E. Margin Loss

To accelerate convergence, we introduce a margin loss in (1), which explicitly penalizes the model if the aggregated cost of generated rollouts, C_{neg}, does not exceed the expert demonstration cost, $C(\xi_{demo})$, by a predefined margin δ.

$$\mathcal{L}_{\mathrm{margin}} = \log\left(1 + \exp\left(-\frac{C_{neg} - C(\xi_{demo}) - \delta}{T}\right)\right), \quad (1)$$

$$\text{where } C_{neg} = \sum_i w_i C(\xi^i_{neg}),\ w_i = \mathrm{softmax}\left(-\frac{C(\xi^i_{neg})}{\tau_{soft}}\right).$$

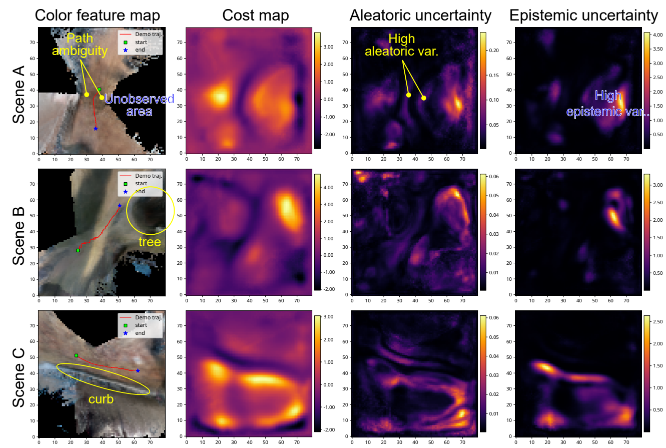

Fig. 3: Results on representative terrain features in the test set.

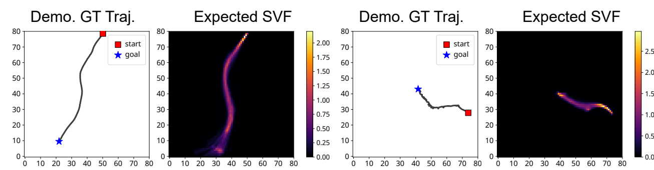

Fig. 4: Comparison between the expert demonstration trajectories and the model's predicted SVF.

III. EXPERIMENT

A. Qualitative Results of Cost and Uncertainty Maps

As shown in Fig. 3, the model accurately assign high costs to physical obstacles like tree (Scene B) and curb (Scene C). Furthermore, the decomposed uncertainties successfully isolate distinct risk sources. Specifically, in Scene A, areas with high path ambiguity yield high aleatoric variance, whereas unobserved regions produce high epistemic variance.

B. Alignment of Expert Trajectories and Expected SVF

As shown in Fig. 4, the expected SVF closely matches the ground truth trajectories, demonstrating successful replication of expert navigation. Our proposed margin loss drives this precise convergence by strictly penalizing deviations from the expert paths.

IV. CONCLUSION

We proposed an uncertainty-aware MEDIRL framework that disentangles aleatoric and epistemic uncertainties via a CVAE and decoder ensemble. By incorporating a margin loss and continuous-state rollouts, the model accurately replicates expert navigation and identifies hazards. Future work will integrate these maps into a downstream path planner.

REFERENCES

[1] S. Baee *et al.*, "MEDIRL: Predicting the visual attention of drivers via maximum entropy deep inverse reinforcement learning," in *Proc. IEEE Intl. Conf. on Comput. Vision*, 2021, pp. 13 178–13 188.

[2] J. Kim, J. Kong, and J. Son, "Conditional variational autoencoder with adversarial learning for end-to-end text-to-speech," in *Proc. Intl. Conf. on Machine Learning*, 2021, pp. 5530–5540.

[3] S. Triest *et al.*, "Learning risk-aware costmaps via inverse reinforcement learning for off-road navigation," *arXiv preprint arXiv:2302.00134*, 2023.