M^2EIT: Multi-Domain Mixture of Experts for Robust Neural Inertial Tracking

Supplementary Material

In this supplementary, we provide more details about baselines and more experimental results, as well as the qualitative analyses.

1. Implementation Details

The number of layers for the temporal decomposition expert and the frequency decomposition expert is set to 2. In the cross-domain translation layer, the number of heads is set to 4. As for parameters within Mamba block, $d\_conv = 2$ and hidden dimension $d\_state = 16$ on all datasets.

2. Baselines

Aside from our proposed M^2EIT, we compare the proposed M^2EIT against the conventional baselines and up-to-date deep models:

- *PDR*: We utilize a stepcounting algorithm [3] to detect foot-steps and move the position along the device heading direction by a predefined distance of 0.67m per step.
- *RIDI*: We employ the original implementation to train distinct models for each device attachment in the RIDI dataset. Conversely, for the remaining datasets, we train a single model for each dataset independently, as the data acquisition process involved mixed attachments [5].
- *RoNIN-ResNet*: RoNIN [1] provides two variants (*i.e. RoNIN-LSTM* and *RoNIN-ResNet*) for regressing planar velocity from the IMU measurements, among which RoNIN-ResNet performs best on the whole testset.
- *TLIO*: In our comparative experiments, we utilized only the neural network components from the publicly available implementation, omitting the EKF module. This configuration is denoted as TLIO (w/o EKF) in the subsequent sections.
- *IMUNet*: This is an extended iteration of RoNIN-ResNet, enhanced with depth-wise and point-wise convolutions to boost the model's real-time performance. We also compared M^2EIT with the two variants (*i.e. Mobilenet* and *EfficientNet*) introduced [6].
- *ResMixer*: Since the official code is not publicly available, we developed our own implementation of ResMixer [2], in which a five-layer mixer layer [4] is used to replace the complex ResNet residual blocks, thereby accelerating inertial odometry inference.

3. Ablation Study

Impact of Mixture of Experts. To gain the insights into our four interaction cells, we conduct ablation studies incrementally. To be more specific, we compared our model M^2EIT with the following variants: 1) SDE: velocity regression using only the spatial decomposition expert; 2) joint representation of the spatial decomposition expert and the temporal decomposition expert;3) our proposed M^2EIT.

Table 1. The results on the impact of Mixture of Experts on the IMUNet dataset.

SDE	TDE	FDE	ATE	RTE
✓	✗	✗	3.14	3.27
✓	✓	✗	2.71	2.83
✓	✓	✓	2.18	2.50

Table 2. Comparison of different wavelet configurations in the M^2EIT framework on RIDI dataset.

Model	ATE	RTE
M^2EIT + Haar	1.65	2.00
M^2EIT + Daubechies	1.70	2.03
M^2EIT + Coiflets	1.66	1.98

As shown in Table 1, progressively incorporating decoupled expert information leads to a steady decline in both ATE and RTE. Notably, in terms of ATE, our approach achieves a 30.6% and 19.6% improvement in localization accuracy compared to SDE and SDE+TDE, respectively. Meanwhile, as we can observe from Figure 2, when the information from the three experts is fully integrated, the generated trajectory becomes closer to the ground truth trajectory. This underscores the critical role of multi-domain information in enhancing inertial representation. Additionally, in Figures 2 and 3, we present the trajectory visualizations of our method on various datasets.

Impact of Frequency Decompositions. Table 2 includes individual expert performance and comparisons with additional frequency decompositions (Haar, Daubechies, Coiflets).

Robustness analysis. Inertial sensors are susceptible to errors, leading to tracking divergence. Therefore, robustness analysis is essential. Based on this, we conduct a series of experiments to observe the model's performance under different noise levels(*i.e.* $[0.01, 0.05, 0.1]$). Figure 1 demon-

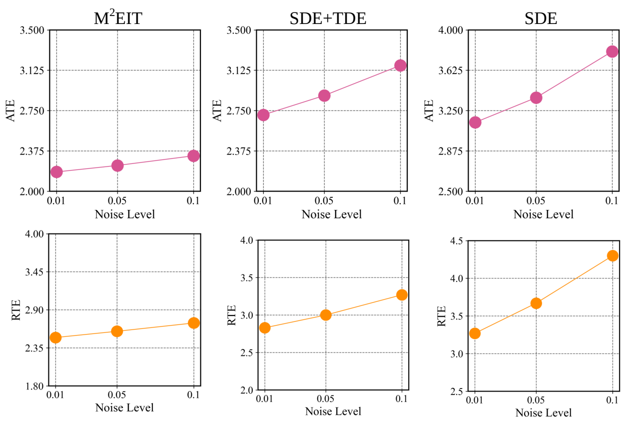


Figure 1. Robustness analysis of M^{2}EIT, SDE+TDE, and SDE on the IMUNet dataset.

strates that the ATE and RTE of M^{2}EIT increase gradually as noise levels rise, whereas the ATE and RTE of SDE and SDE+TDE escalate significantly, indicating their heightened sensitivity to noise. Evidently, incorporating more domain-specific information enhances the model's robustness.

References

[1] Sachini Herath, Hang Yan, and Yasutaka Furukawa. Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods. In *2020 IEEE international conference on robotics and automation (ICRA)*, pages 3146–3152. IEEE, 2020. 1

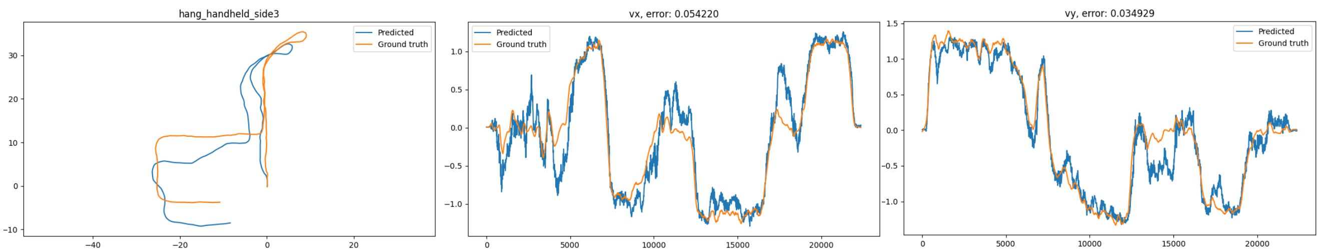
[2] Rucong Lai, Yong Tian, Jindong Tian, Jie Wang, Ning Li, and Yue Jiang. Resmixer: A lightweight residual mixer deep inertial odometry for indoor positioning. *IEEE Sensors Journal*, 2024. 1

[3] Qinglin Tian, Zoran Salcic, I Kevin, Kai Wang, and Yun Pan. An enhanced pedestrian dead reckoning approach for pedestrian tracking using smartphones. In *2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pages 1–6. IEEE, 2015. 1

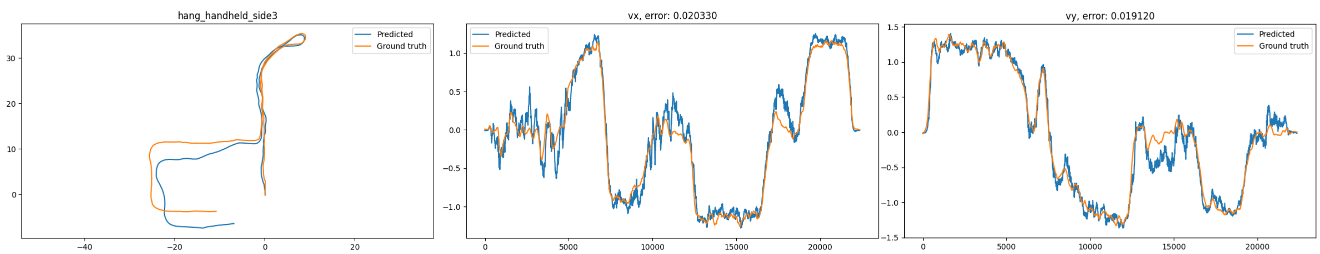
[4] Ilya O Tolstikhin, Neil Houlsby, Alexander Kolesnikov, Lucas Beyer, Xiaohua Zhai, Thomas Unterthiner, Jessica Yung, Andreas Steiner, Daniel Keysers, Jakob Uszkoreit, et al. Mlp-mixer: An all-mlp architecture for vision. *Advances in neural information processing systems*, 34:24261–24272, 2021. 1

[5] Hang Yan, Qi Shan, and Yasutaka Furukawa. Ridi: Robust imu double integration. In *Proceedings of the European conference on computer vision (ECCV)*, pages 621–636, 2018. 1

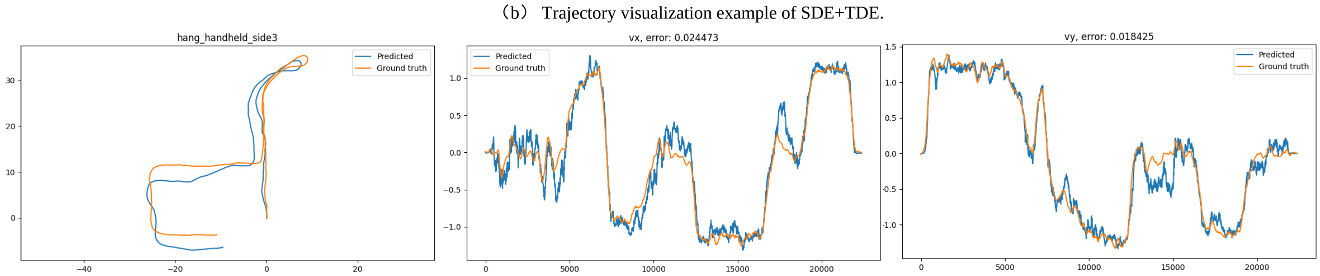
[6] Behnam Zeinali, Hadi Zanddizari, and Morris J Chang. Imunet: Efficient regression architecture for inertial imu navigation and positioning. *IEEE Transactions on Instrumentation and Measurement*, 2024. 1



（a） Trajectory visualization example of SDE.

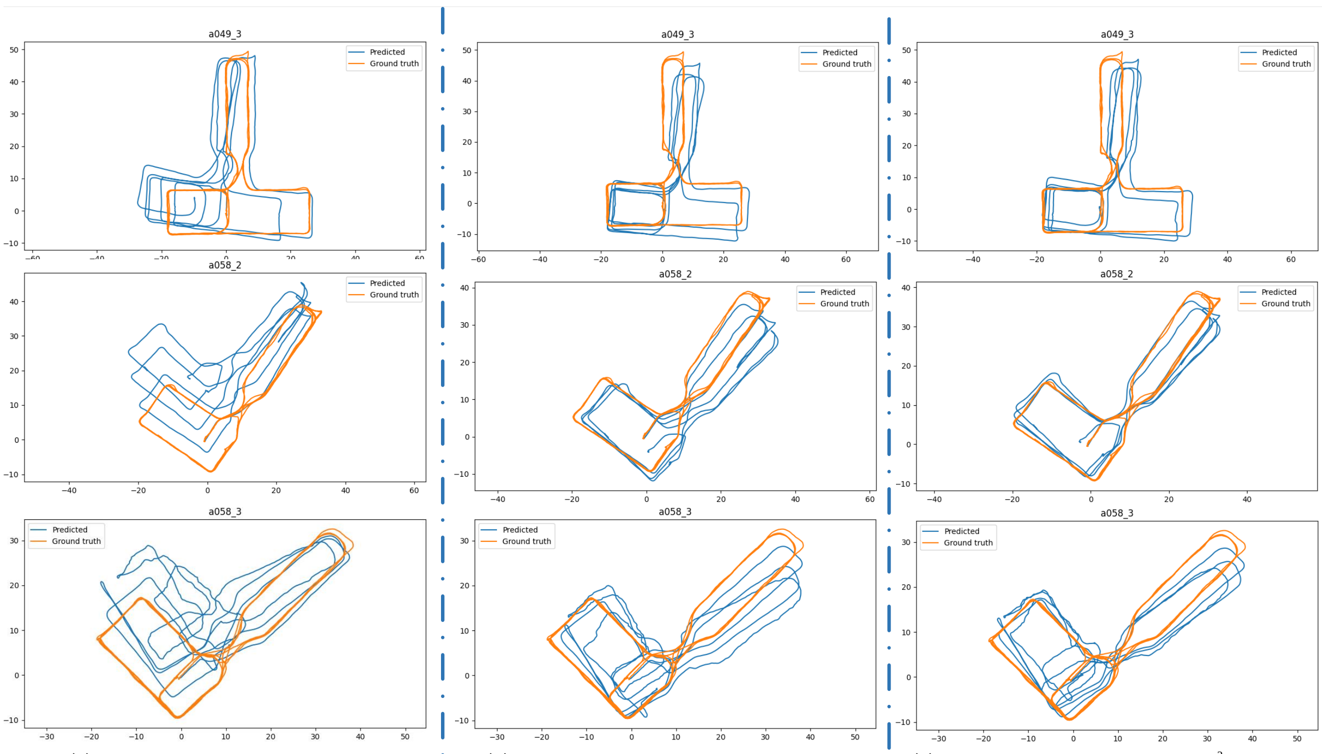


（b） Trajectory visualization example of SDE+TDE.



（c） Trajectory visualization example of M^2EIT.

Figure 2. Trajectory visualization example of SDE, SDE+TDE, and M^2EIT on the RIDI dataset.



(a) Trajectory visualization example of SDE.　(b) Trajectory visualization example of SDE+TDE.　(c) Trajectory visualization example of M^{2}EIT.

Figure 3. Trajectory visualization example of SDE, SDE+TDE, and M^{2}EIT on the RONIN dataset.