RIO: Rotation-equivariance supervised learning of robust inertial odometry Supplementary Material

Xiya Cao[1*] Caifa Zhou[1] Dandan Zeng Yongliang Wang
Riemann lab, 2012 Laboratories, Huawei Technologies Co.Ltd

The supplementary materials provides details of the pseudo-code of adaptive TTT, the description of IPS dataset and additional results of RIO.

1. Pseudo-code of adaptive TTT

The pseudo-code of adaptive TTT is as follows:

Algorithm 1 Adaptive TTT

Initialize model $F(.|\theta)$ with pre-trained θ^* as θ_0
Initialize ensemble models $\{F(.|\theta_i)\}_{i=1}^{M}$
for X_t arrive in a batch at time t **do** ▷ 128 samples
  Compute ensemble outputs $\{\vec{v}_i^t = F(X_t|\theta_i)\}_{i=1}^{M}$
  Compute average output $\vec{v}_*^t = \frac{1}{M}\sum_i \vec{v}_i^t$
  Compute variance $\sigma_{v_t}^2 = \frac{1}{M}\sum_i (\vec{v}_i^t - \vec{v}_*^t)^2$
  if $min(\sigma_{v_t}^2) < 1e-4$ **then**
    Restore model $F(.|\theta)$ with pre-trained θ^* as θ_t
  else if $mean(\sigma_{v_t}^2) < 0.04$ **then**
    Keep model parameters unchanged $\theta_t = \theta_{t-1}$
  else
    for $iter = 1$ to N_{TTT} **do**
      Select k angles $\{\phi_i\}_{i=1}^{k}$
      Rotate X_t and get $\{X_t^{\phi_i}\}_{i=1}^{k}$
      $\vec{v}_t = F(X_t|\theta_{t-1})$ ▷ original
      Rotate $\vec{v}_t$ and get $\{\vec{v}_t^{\phi_i}\}_{i=1}^{k}$
      $\{\vec{v}_t^{c_i} = F(X_t^{\phi_i}|\theta_{t-1})\}_{i=1}^{k}$ ▷ conjugate
      $l_{ssl_t} = \sum_i \mathcal{L}(\vec{v}_t^{\phi_i}, \vec{v}_t^{c_i})$ ▷ loss
      Update θ using Adam
    end for
  end if
  Compute predictions $\vec{v}_t = F(X_t|\theta_t)$
end for

2. IPS dataset

In order to evaluate the robustness of proposed inertial odometry against the type, placement, and handheld poses of mobile devices, we have collected 155 trajectories using a custom-developed application for logging the IMU measurements. The summary of IPS dataset is in Tab. 1 . As for the ground truth, we use the dual foot-mounted IMU using AW70[1]. The synchronization between mobile device and foot-mounted IMU is via Bluetooth connection. The IMU data from the AW70 and from mobile device are logged on the mobile device side. So that all the timestamps are the same as the mobile device.

Table 1. IPS database summary

Scenarios	Time(s)	Distance(m)	Count
Reading portrait	1726	1637	9
Reading landscape	2531	2446	13
Calling	1965	1818	10
Bag	1742	1649	9
Front pocket	1756	1687	9
Back pocket	2088	2019	11
Photo portrait	2090	2089	11
Photo landscape	2397	2289	12
Swing arm	1874	1848	10
Combine scenes	13857	12166	61
Total	32027	29648	155

3. Results of uncertainty estimation

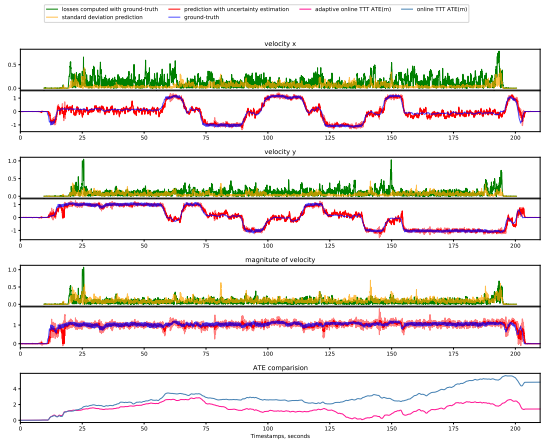
Trajectories from IPS, OXIOD and RIDI databases are relative short trajectories and do not have stationary status. However, uncertainty estimation still capture some error peaks as show in Fig. 1. Adaptive TTT strategy with the uncertainty estimation leads to small ATE over time.

4. Results of impact of size of training data and TTT iterations

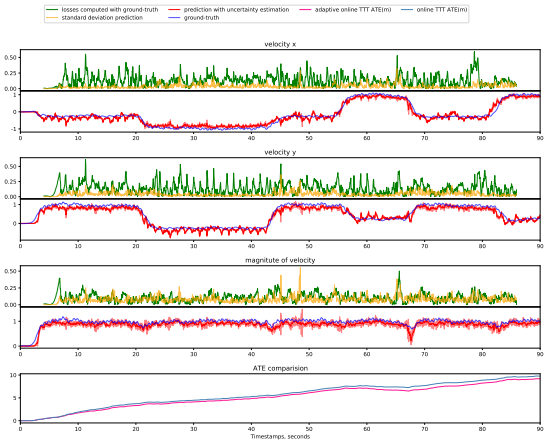
We present influence of training data size and updating iterations on model performance with RoNIN, OXIOD and RIDI databases in Fig. 2. They yield consistent conclusion as with IPS database that i) *J-ResNet-TTT* with 30%

[*1] denotes equal contribution.

[1]Details about the AW70 can found in https://consumer.huawei.com/levant/support/wearables/band3e/.



(a) Example uncertainty estimation on IPS



(b) Example uncertainty estimation on OXIOD

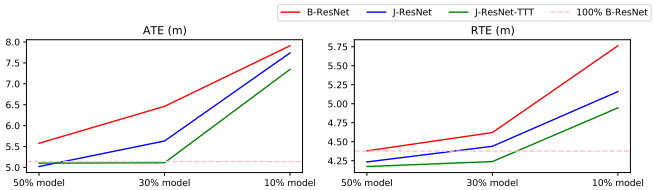
(c) Example uncertainty estimation on RIDI

Figure 1. Uncertainty estimations for RoNIN, OXIOD and RIDI Databases.

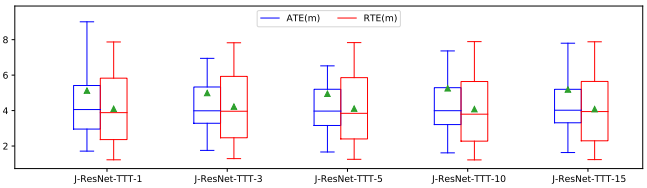
training database achieves on par performance wirh 100% B-ResNet, ii) more iterations result in model performance improvement when number of iterations less than 5.

5. Visualization of trajectories

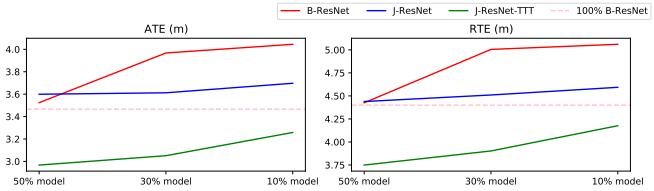
We select 9 trajectories from RoNIN, OXIOD and RIDI databases and visualize estimated results of the same model with or w/o the proposed TTT strategy in Fig. 3. It shows that for trajectory Fig. 3e that model already works well, the proposed TTT strategy leads to little improvement. It yields better performance for trajectories Fig. 3a, Fig. 3h, Fig. 3i. For trajectories Fig. 3b, Fig. 3c, Fig. 3d, Fig. 3f, Fig. 3g that model has not adapted to, TTT strategy improves a lot.



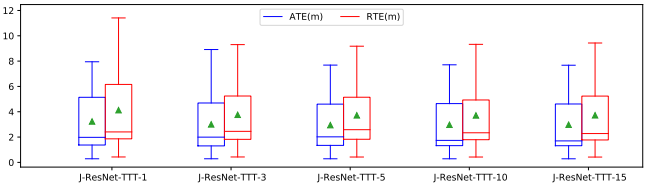
(a) Impact of the number of TTT iterations on ATE and RTE on RoNIN



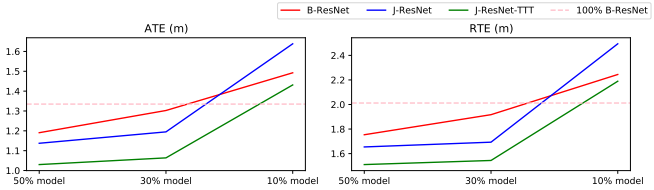
(b) Impact of the number of TTT iterations on ATE and RTE on RoNIN



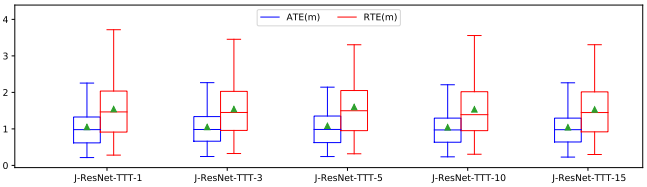
(c) Impact of the number of TTT iterations on ATE and RTE on OXIOD



(d) Impact of the number of TTT iterations on ATE and RTE on OXIOD

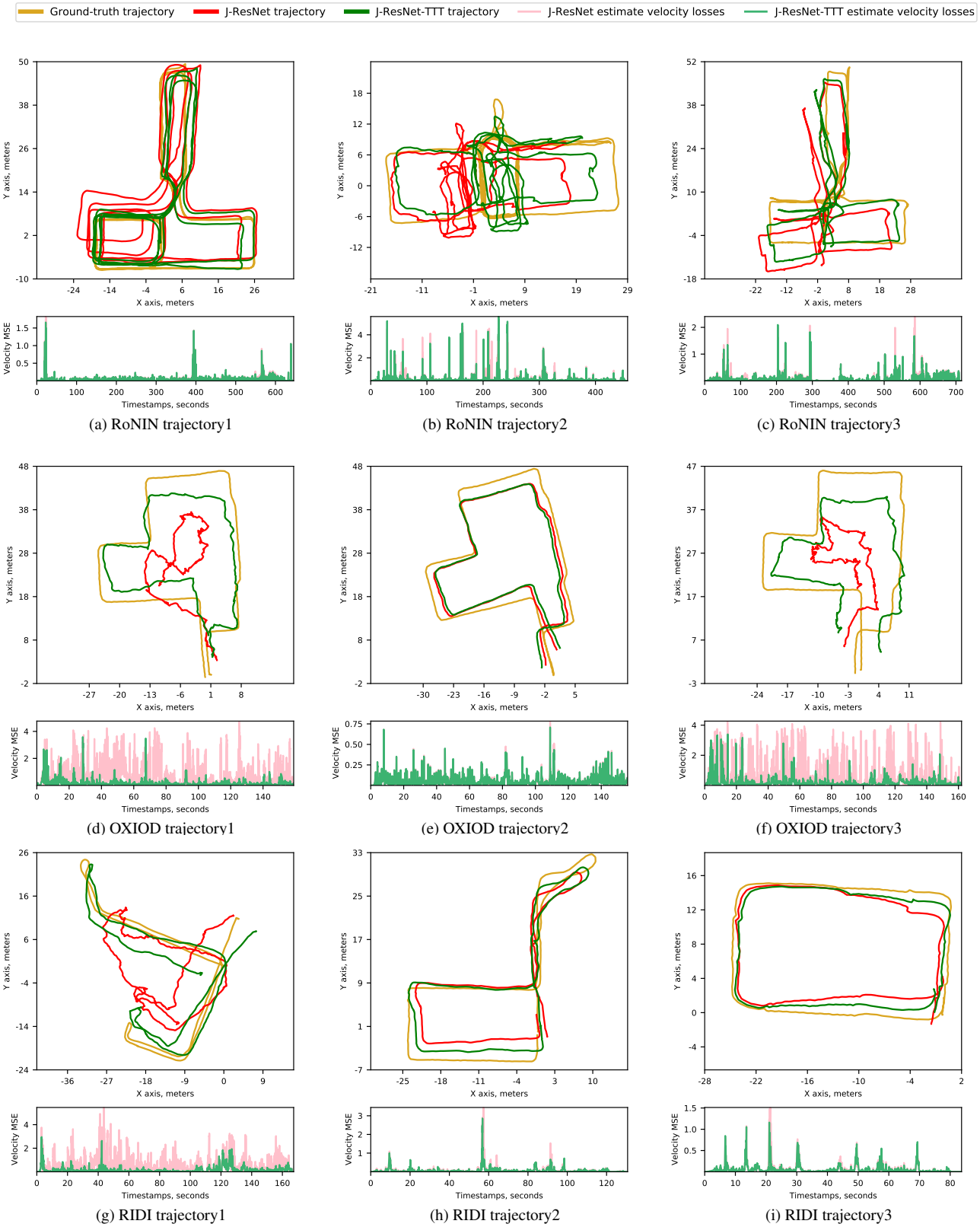


(e) Impact of the number of TTT iterations on ATE and RTE on RIDI



(f) Impact of the number of TTT iterations on ATE and RTE on RIDI

Figure 2. Summary of the size of training data and TTT iterations on various datasets



(a) RoNIN trajectory1 (b) RoNIN trajectory2 (c) RoNIN trajectory3

(d) OXIOD trajectory1 (e) OXIOD trajectory2 (f) OXIOD trajectory3

(g) RIDI trajectory1 (h) RIDI trajectory2 (i) RIDI trajectory3

Figure 3. Selected visualizations. We select 3 examples from each open sourced dataset and visualize trajectories of ground-truth, J-ResNet and J-ResNet-TTT, along with velocity losses comparison of J-ResNet and J-ResNet-TTT.