iMoT: Inertial Motion Transformer for Inertial Navigation

Son Minh Nguyen Duc Viet Le Paul Havinga

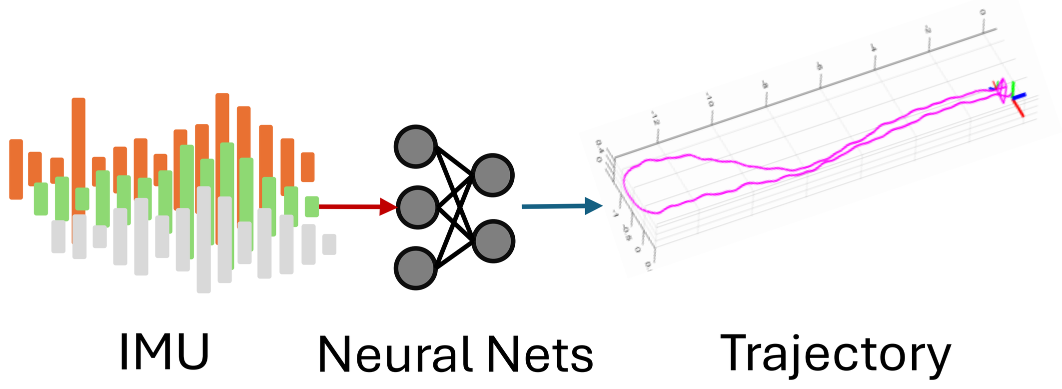
Department of Computer Science, University of Twente, Enschede, The Netherlands

UNIVERSITY OF TWENTE.

MOSAIC enhancement of MicrOfluidic Sensing with Artificial IntelligenCe

AAAI-25 / IAAI-25 / EAAI-25
FEBRUARY 25 – MARCH 4, 2025 | PHILADELPHIA, USA

Introduction



- **Inertial Measurement Unit (IMU)** signals, constituted by **Acceleration**, **Angular Velocity**, and **Geomagnetic**, primarily represent changes in motion and rotation.
- **Inertial Odometry (IO) Navigation** aims to reconstruct the traveled trajectory of a subject through recorded IMU measurements.

Motivation

Modality advantage

- **Resilience:** Compared to common passive (i.e., reflection from beaming) modalities, e.g., visual, acoustic, and radar information, IMU signals can be independently measured, and are much less vulnerable to environmental dynamics.
- **Efficiency:** Requiring less energy to operate.
- **Omnipresence:** Widely integrated in most mobile phones.

SoTA limitations

- **Motion Uncertainty:** Although the human walking style is symmetric and repeated, each individual retains a subtly distinct gait.
- **Unimodal Exploitation:** Despite the difference in expression, acceleration, and angular velocity are often exploited in a unimodal learning fashion.

Approach

Inertial Motion Transformer (iMoT) is proposed to ease these above issues.

- **Encoder** to aggregate context features from motion and rotation information over T measurements of one second.
- **Decoder** responsible for exploiting cross-modal information to represent uncertainties in motion by manipulating a set of learnable query motion particles.

Architecture

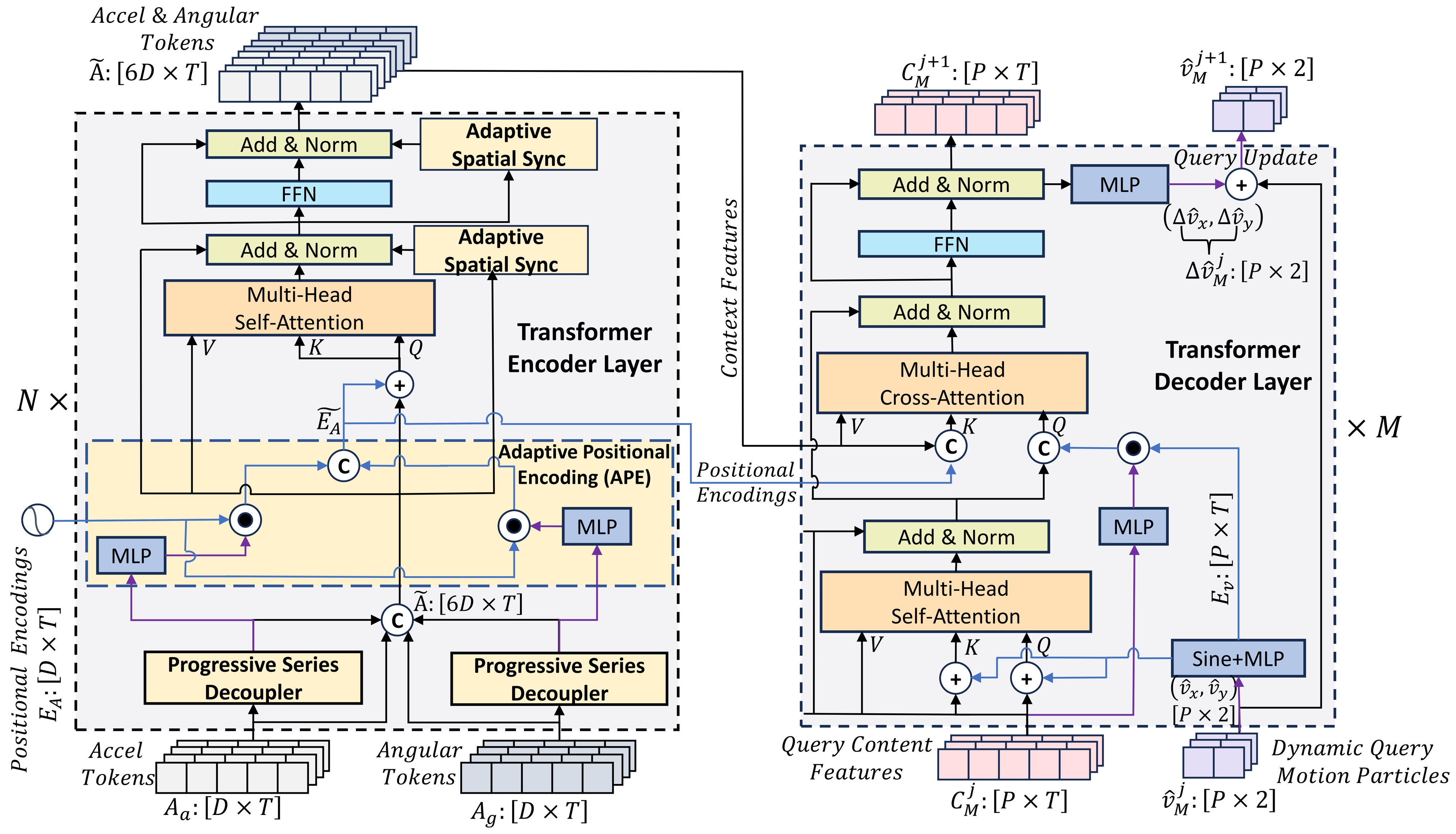


Figure 1. Inertial Motion Transformer (iMoT).

Trajectory Visualization

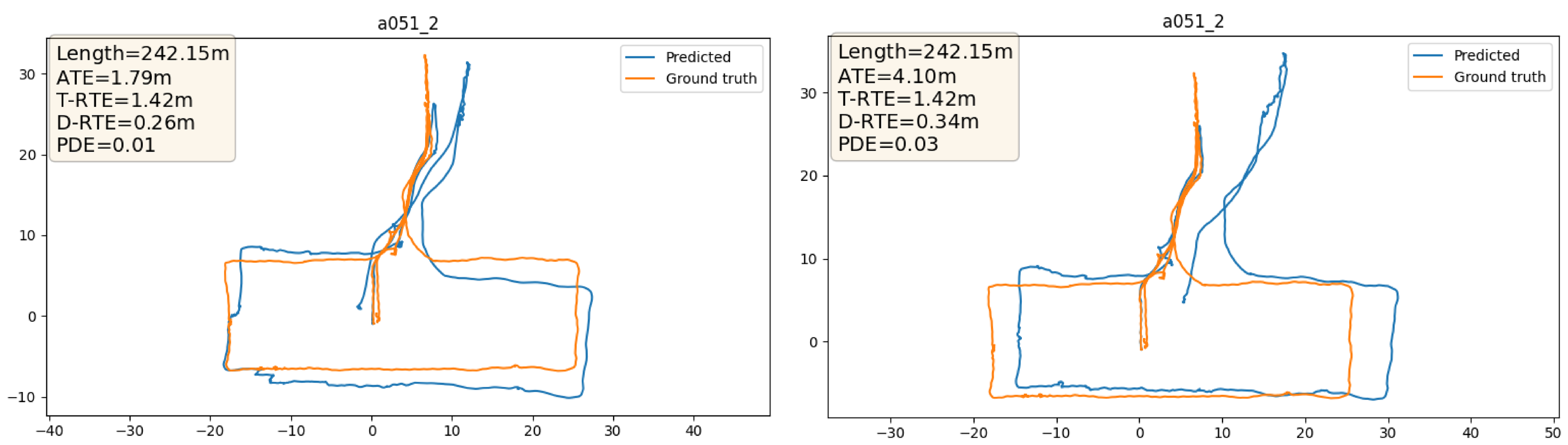


Figure 2. Reconstructed trajectories of unseen subjects for iMoT, CTIN on the RoNIN.

Priciple Components

Encoder

- **Progressive Series Decoupler** to decompose IMU signals into two more interpretable components, *e.g.*, seasonal signals and trend-cycle signals, highlighting critical motion events such as half-turns, U-turns, and periods of stillness better.
- **Adaptive Positional Encoding** allows for tailoring the positional encoding to specific characteristics of motion and rotation data.
- **Adaptive Spatial Sync Encoding** incorporates spatial interactions across sensor channels at each time step.

Decoder

- **Query Motion Particles** are iteratively refined to account for motion uncertainties among individuals in the form of learnable positional embeddings.
- **Query Content Features** store cross-modal information and are referred to adjust the motion particles for the subsequent decoding steps.
- **Dynamic Scoring Mechanism** adaptively synthesizes all particles at the last layer into desired velocity segments.

Analysis

SoTA Comparison

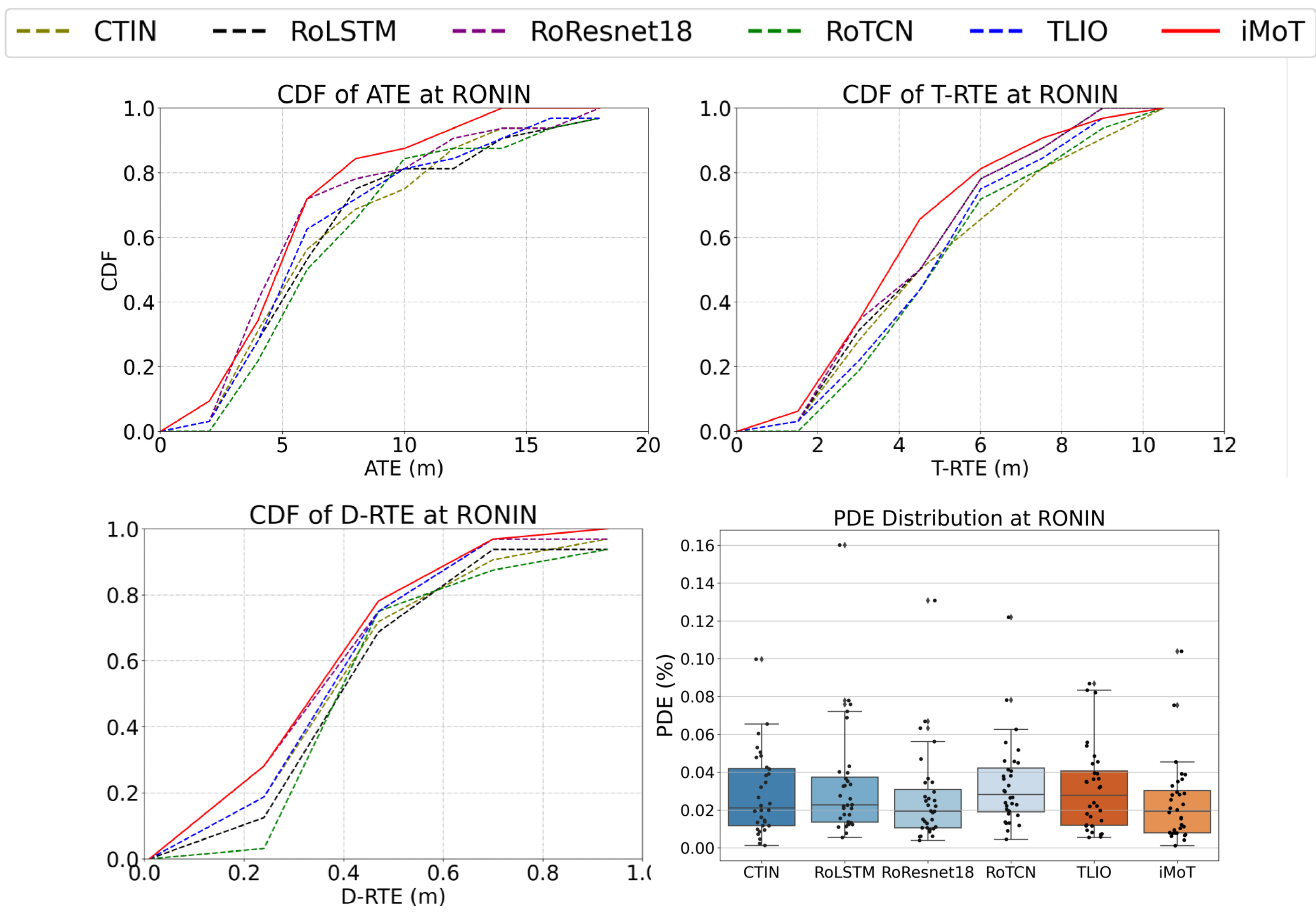


Figure 3. Cumulative Error Distributions (CDF) with three types of metric types, and boxplot of PDE on RoNIN dataset.

Benchmark Datasets:
OxIOD, IDOL, RIDI, and RoNIN with dynamic attachments of different types of IMU-recording devices.

Metrics:

- Absolute Trajectory Error (ATE) (m).
- Time-Relative Trajectory Error (T-RTE) (m).
- Distance-Relative Trajectory Error (D-RTE) (m).
- Position Drift Error (PDE) (%).

Conclusion

- Progressive Series Decoupler facilitates the absorption of complex IMU signals.
- Cross-modal Exploitation of IMU enables the all-round enhancement of estimation quality.
- Manipulation of learnable motion query particle set can represent motion uncertainties more effectively .