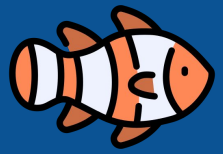
ICLR 2026



NeMo-map: Neural Implicit Flow Fields for Spatio-Temporal Motion Mapping

Robot Navigation & Perception Lab

Authors:

Yufei Zhu[1], Shih-Min Yang[1],
Andrey Rudenko[2], Tomasz P. Kucner[3],
Achim J. Lilienthal[1,2], Martin Magnusson[1]

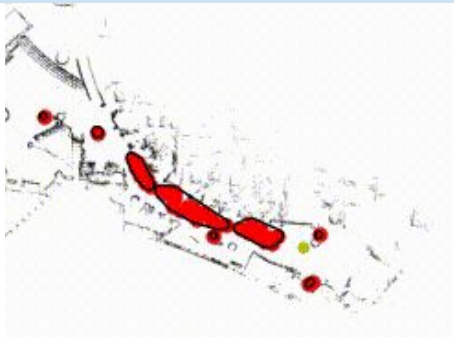
ÖREBRO UNIVERSITY

Aalto University
School of Science and Technology

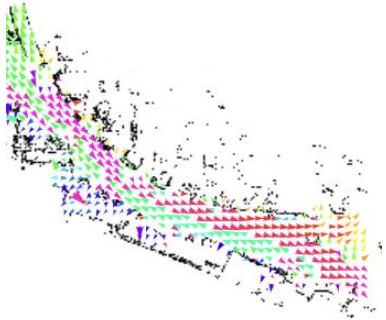
1. AASS RNP lab, Örebro University, Sweden
2. TU Munich, Germany.
3. Aalto University, Finland.

NeMo-map:

A continuous **spatio-temporal Map** of Dynamics using implicit neural representations, achieving more accurate and efficient human motion pattern representation.



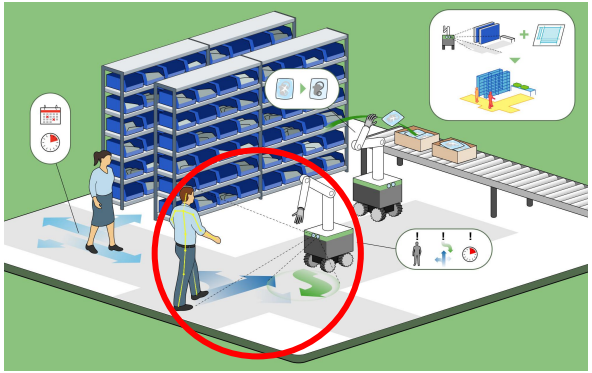
From local human **motion observations**



To spatiotemporal motion **patterns**

Why Spatiotemporal Flow Modeling?

Challenge: How can autonomous systems learn and exploit motion dynamics to act proactively rather than reactively?

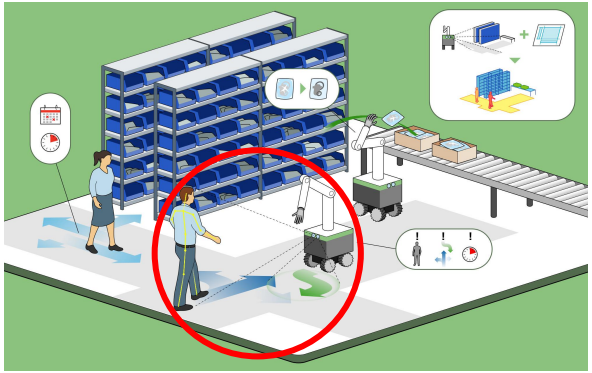


Robots operate in **dynamic**, **uncertain** environments.

Why Spatiotemporal Flow Modeling?

Challenge: How can autonomous systems learn and exploit motion dynamics to act proactively rather than reactively?

- Maps of dynamics (MoDs): provide **motion patterns as a map**, which encode spatiotemporal representations of patterns of dynamics within the environment.



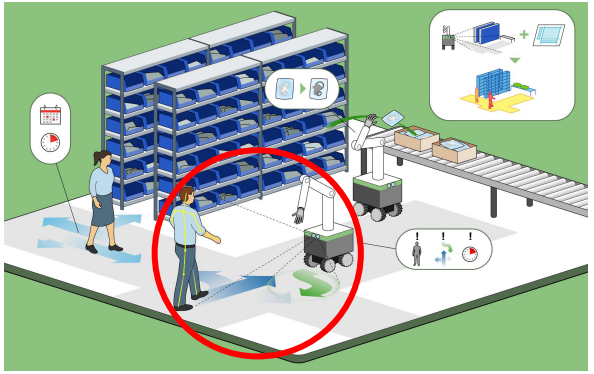
Robots operate in **dynamic, uncertain** environments.

Why Spatiotemporal Flow Modeling?

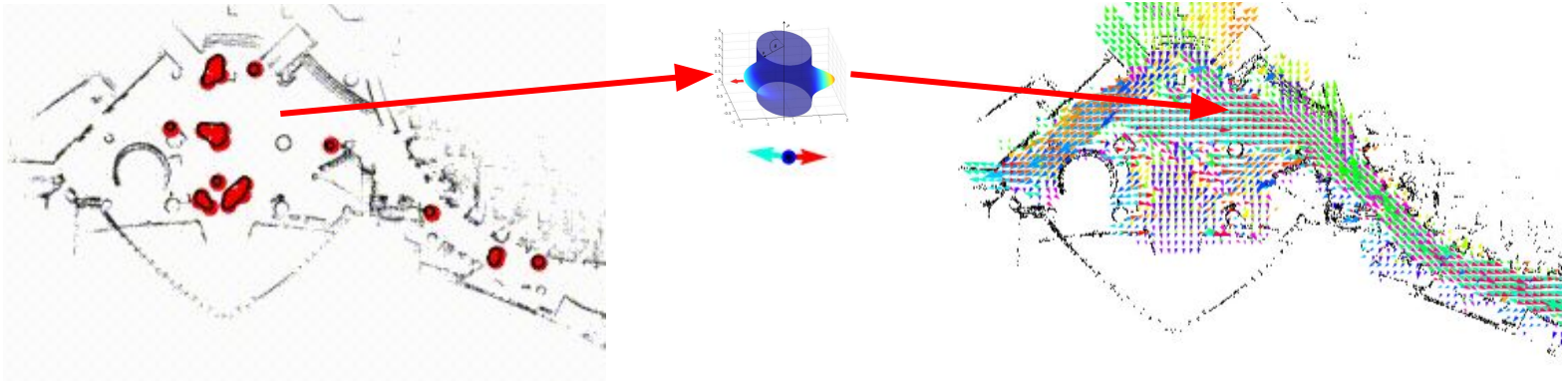
Challenge: How can autonomous systems learn and exploit motion dynamics to act proactively rather than reactively?

- Maps of dynamics (MoDs): provide **motion patterns as a map**, which encode spatiotemporal representations of patterns of dynamics within the environment.

(time, location) → Motion distribution



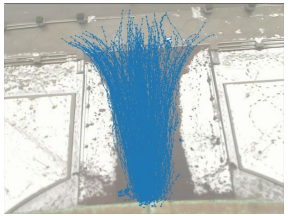
Robots operate in **dynamic, uncertain** environments.



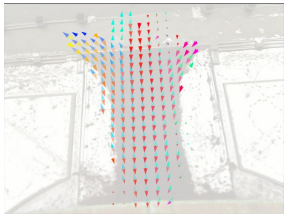
Maps of dynamics

Maps of dynamics (**MoDs**) encode motion patterns that:

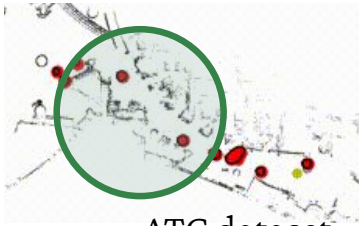
- multimodal
- structured by geometry & physics
- change over space and time



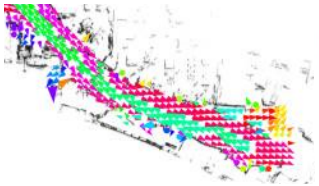
ETH dataset



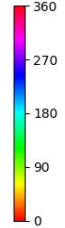
Motion Patterns



ATC dataset



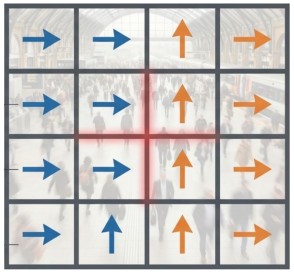
Left hand traffic

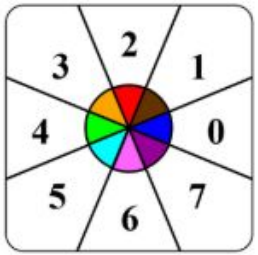


Limitations of Existing MoDs

Existing methods require **discretization**:

- CLiFF-map [1]: fixed grid resolution, per-cell modeling (~30 hours on ATC)
- STeF-map [2]: grid-based, discretizes orientation into 8 bins, no speed modeling





Limitations:

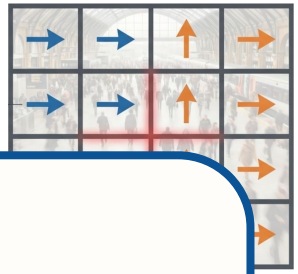
- Introduces information loss
- Reduces accuracy and flexibility
- Complicates tuning across different environments.

[1] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts and A. J. Lilienthal, "Enabling Flow Awareness for Mobile Robots in Partially Observable Environments," in IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 1093-1100, April 2017, doi: 10.1109/LRA.2017.2660060.
[2] S. Molina, G. Cielniak and T. Duckett, "Robotic Exploration for Learning Human Motion Patterns," in IEEE Transactions on Robotics, vol. 38, no. 2, pp. 1304-1318, April 2022, doi: 10.1109/TRO.2021.3101358.

Limitations of Existing MoDs

Existing methods require **discretization**:

- CLiFF-map [1]: fixed grid resolution, per-cell modeling (~30 ho
- STeF-n bins, n



Limitation

- Introdu
- Reduce
- Complicates tuning across different environments.

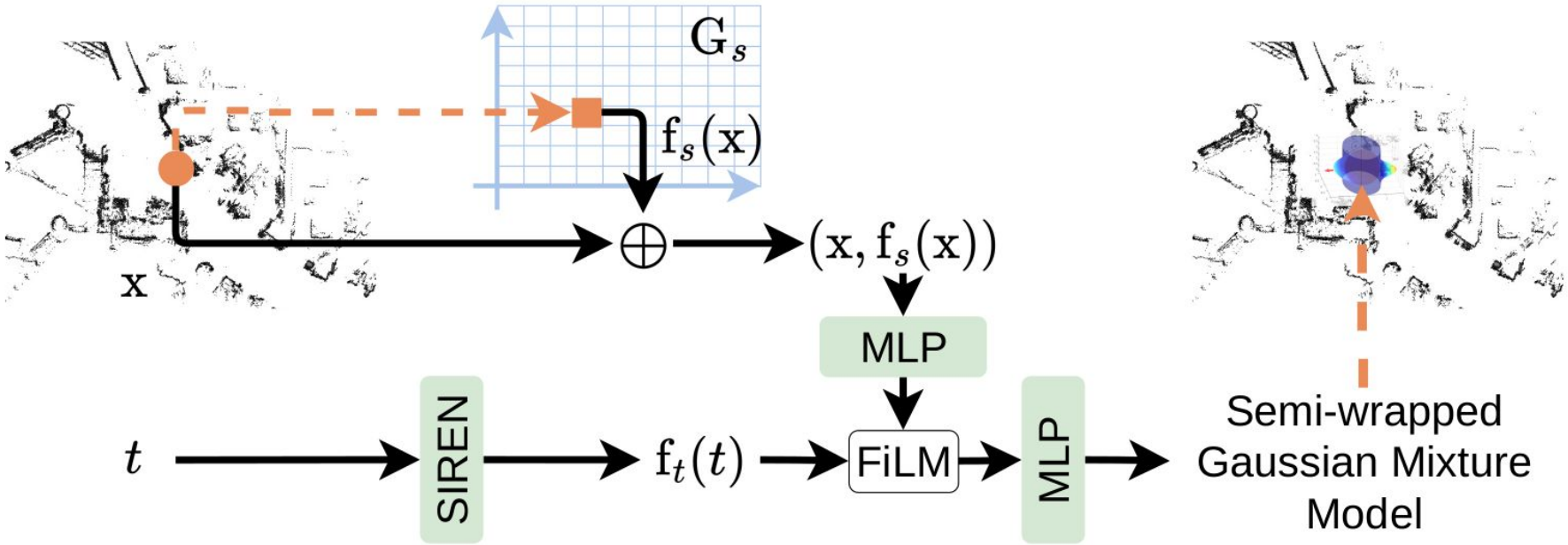
Our goal:
an accurate, continuous representation of human motion dynamics over space and time.

[1] T. P. Kucner et al., "Enabling Flow Awareness for Mobile Robots in Partially Observable Environments," in IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 1093-1100, April 2017, doi: 10.1109/LRA.2017.2660060.

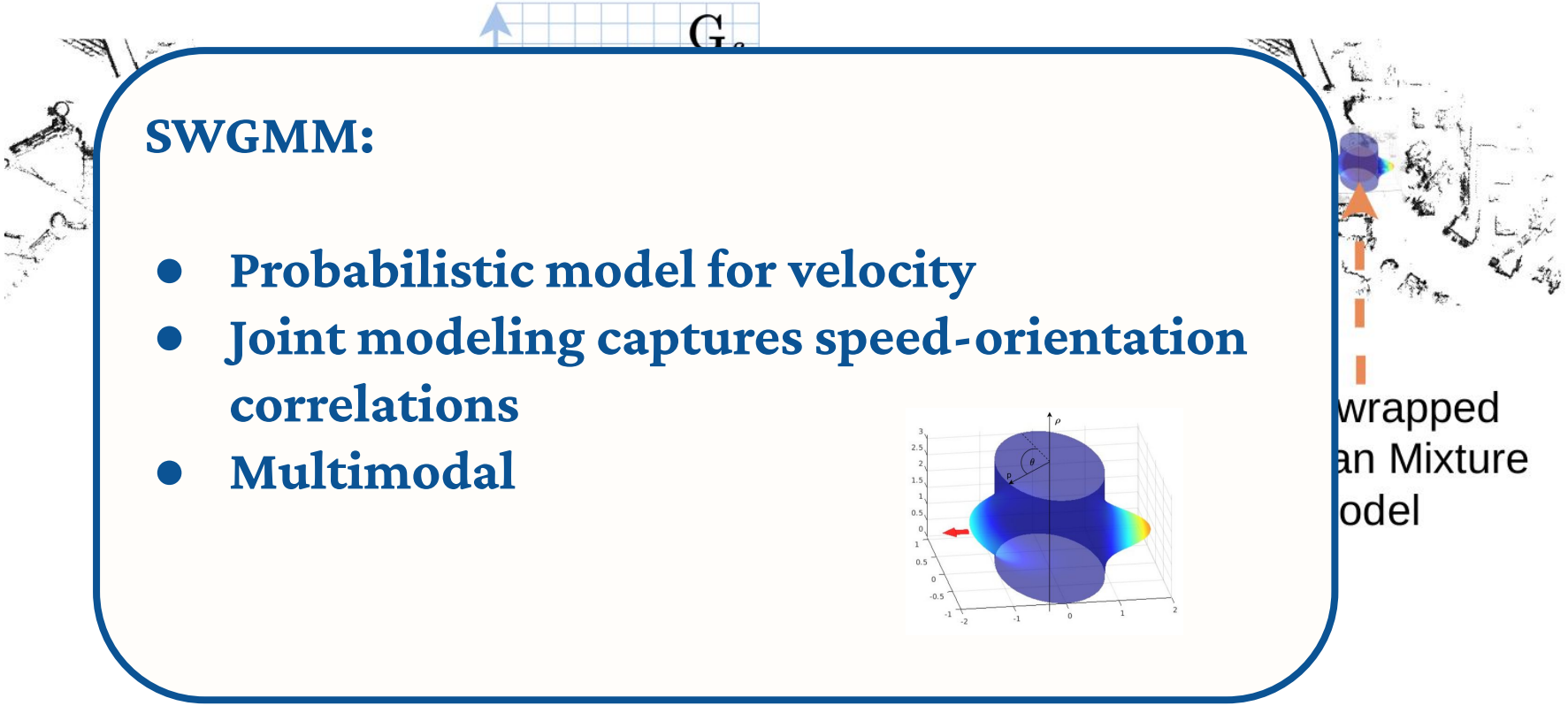
[2] S. Molina et al., "Robotic Exploration for Learning Human Motion Patterns," in IEEE Transactions on Robotics, vol. 38, no. 2, pp. 1304-1318, April 2022, doi: 10.1109/TRO.2021.3101358.

NeMo-map: Method Overview

Learn a continuous function Φ that maps any $(\mathbf{x}, t)$ to a motion distribution



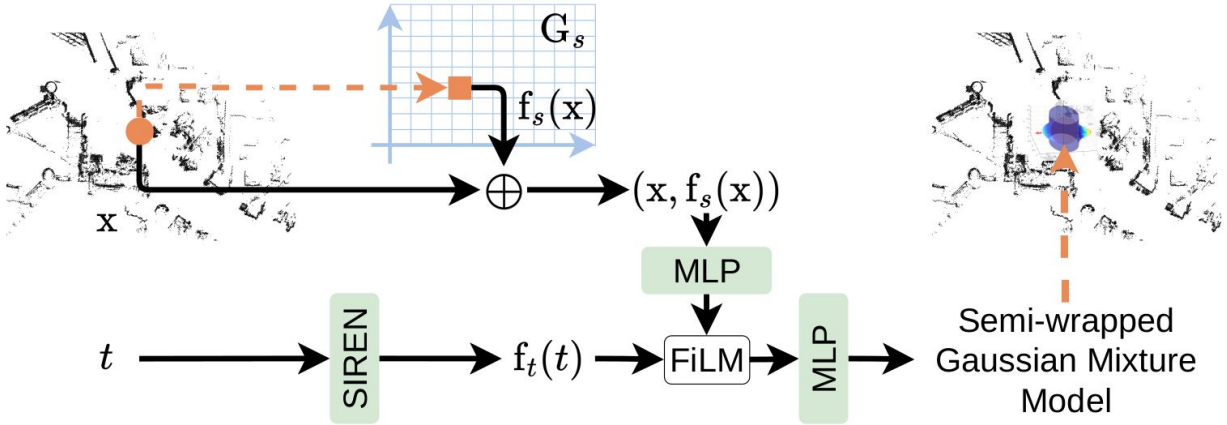
NeMo-map: Method Overview



SWGMM:

- **Probabilistic model for velocity**
- **Joint modeling captures speed-orientation correlations**
- **Multimodal**

NeMo-map: Method Overview



Spatial features:

- Learnable feature grid
- Continuous in space, captures local motion variations

Temporal encoding:

- SIREN [1] network (periodic activations)
- Continuous in time

Training:

- Minimize negative log-likelihood (**NLL**) of observed velocities

[1] Sitzmann et al., "Implicit Neural Representations with Periodic Activation Functions," NeurIPS 2020

Results: Representation Accuracy

Evaluate on two **real-world** datasets:

- ATC: indoor shopping mall, multi-day coverage, temporal variation
- ETH/UCY: outdoor scenes, short sequences, multiple environments

Method	NLL↓	Reduction vs Ours	95% CI
Ours	**0.775 ± 2.052**	–	–
Online CLiFF	1.527 ± 4.156	+0.752	[0.749, 0.755]
CLiFF-map	1.964 ± 4.953	+1.189	[1.185, 1.192]
STeF-map	5.576 ± 9.314	+4.801	[4.794, 4.809]

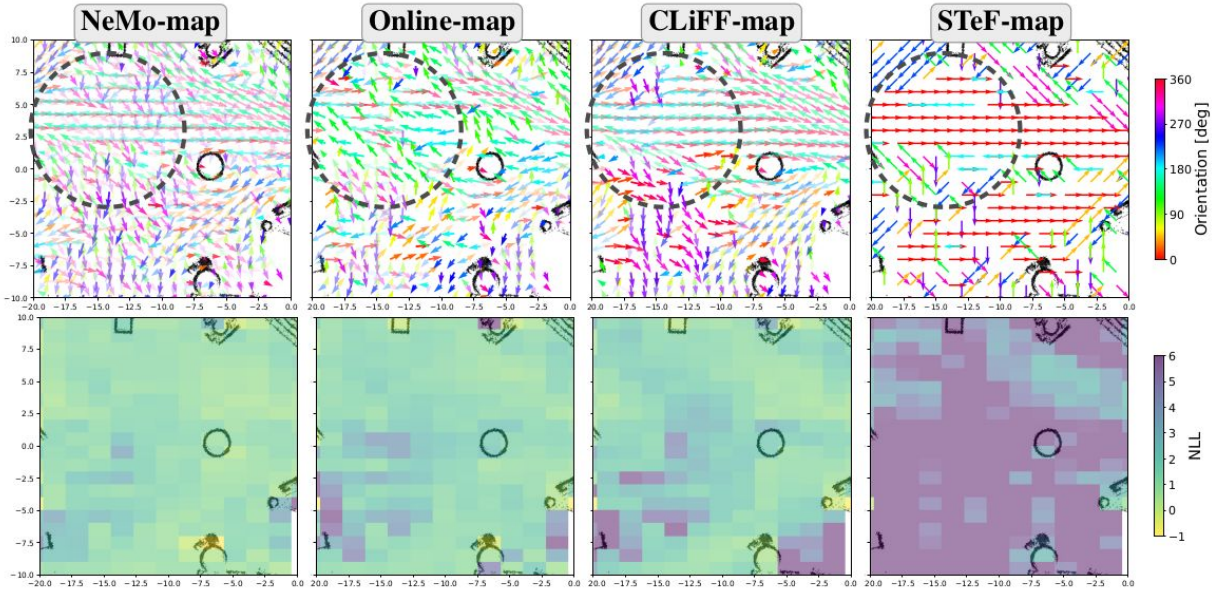
TABLE I: Accuracy on ATC using average NLL (↓). Mean ± std.

Method	ETH	HOTEL	UNIV	ZARA
Ours	**-0.384**	**-0.838**	**0.404**	**-0.342**
CLiFF-map	0.112	0.701	0.518	0.068
Online CLiFF	0.086	1.241	0.577	0.186
STeF-map	2.315	3.349	10.932	2.784

TABLE II: Accuracy on ETH/UCY using NLL (↓).

Accuracy evaluation:
Negative Log-Likelihood (NLL) of observed velocities.
Lower NLL indicates a more accurate representation.

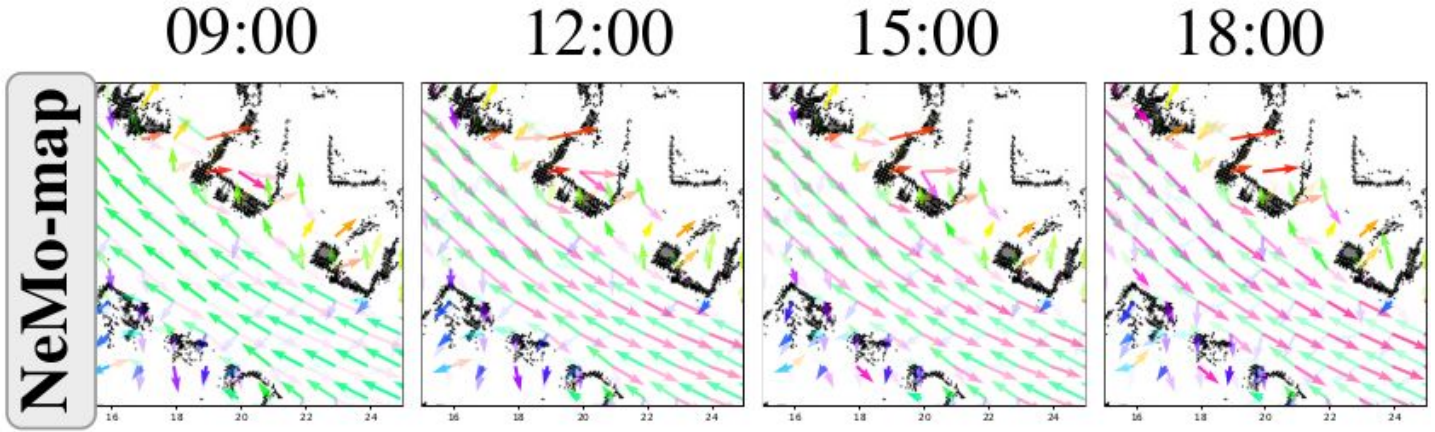
Examples: Spatial Multimodality



Central area: horizontal flow intersects with vertical flow

- NeMo-map preserves both crossing flows continuously
- Baselines lose the vertical flow in the middle section
- Bottom row: NeMo-map achieves lower NLL (lighter = better)

Examples: Temporal Variation

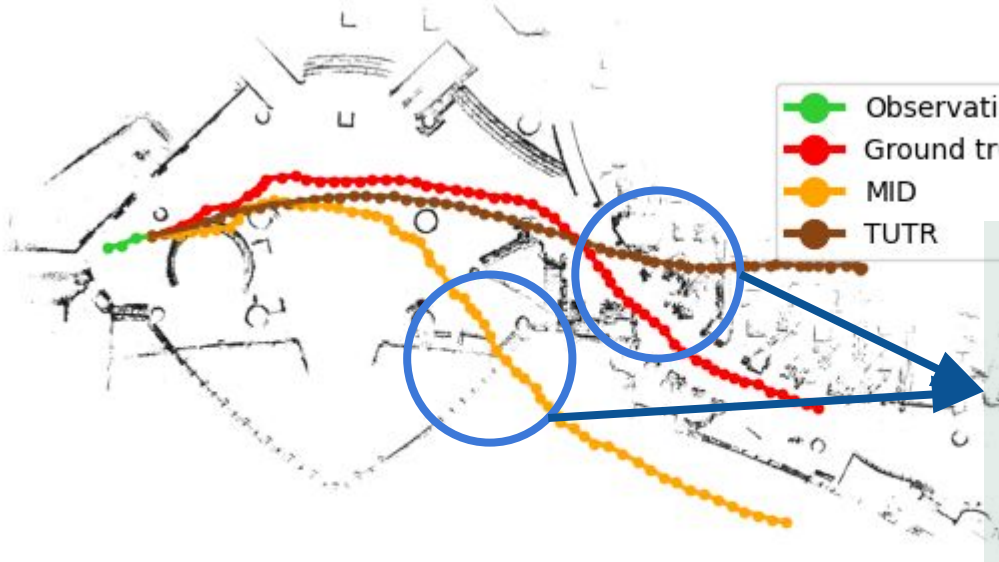


East corridor: dominant flow direction shifts throughout the day

- NeMo-map captures how the dominant flow shifts from morning to evening in a single continuous model.

Downstream Task Evaluation

Evaluate Maps of Dynamics in **long-term human motion prediction task**



Without explicit map guidance, baseline predictors (MID [1], TUTR [2]) produce unfeasible trajectories — e.g., crossing through walls.

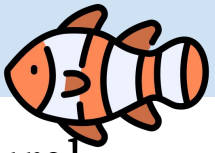
[1] Gu et al., "Stochastic Trajectory Prediction via Motion Indeterminacy Diffusion," CVPR 2022
[2] Shi et al., "Trajectory Unified Transformer for Pedestrian Trajectory Prediction," ICCV 2023

Downstream Task Evaluation

Evaluate Maps of Dynamics in **long-term human motion trajectory task**

With NeMo-map: predictions follow realistic flows, implicitly respect the complex topology of the environment

Summary



- NeMo-map: A continuous spatio-temporal MoD using implicit neural representations
- Maps any (x, y, t) to multimodal flow distribution (SWGMM)
- Best representation accuracy across ATC and ETH/UCY datasets
- Practical and efficiency: ~19 min training (vs ~30h CLiFF-map)
- Improve performance in downstream task: long-term human motion prediction

Future work:

- Online updates for evolving environments
- Handling sharp spatial discontinuities (walls, doors)

Code is available:

https://github.com/test-bai-cpu/nemo-map