

Inferring Terrain Stability from Proprioceptive Data of a Quadruped Robot

Jean-Michel Fortin, Daniele De Martini, François Pomerleau, Philippe Giguère



UNIVERSITÉ
LAVAL



UNIVERSITY OF
OXFORD

Overview

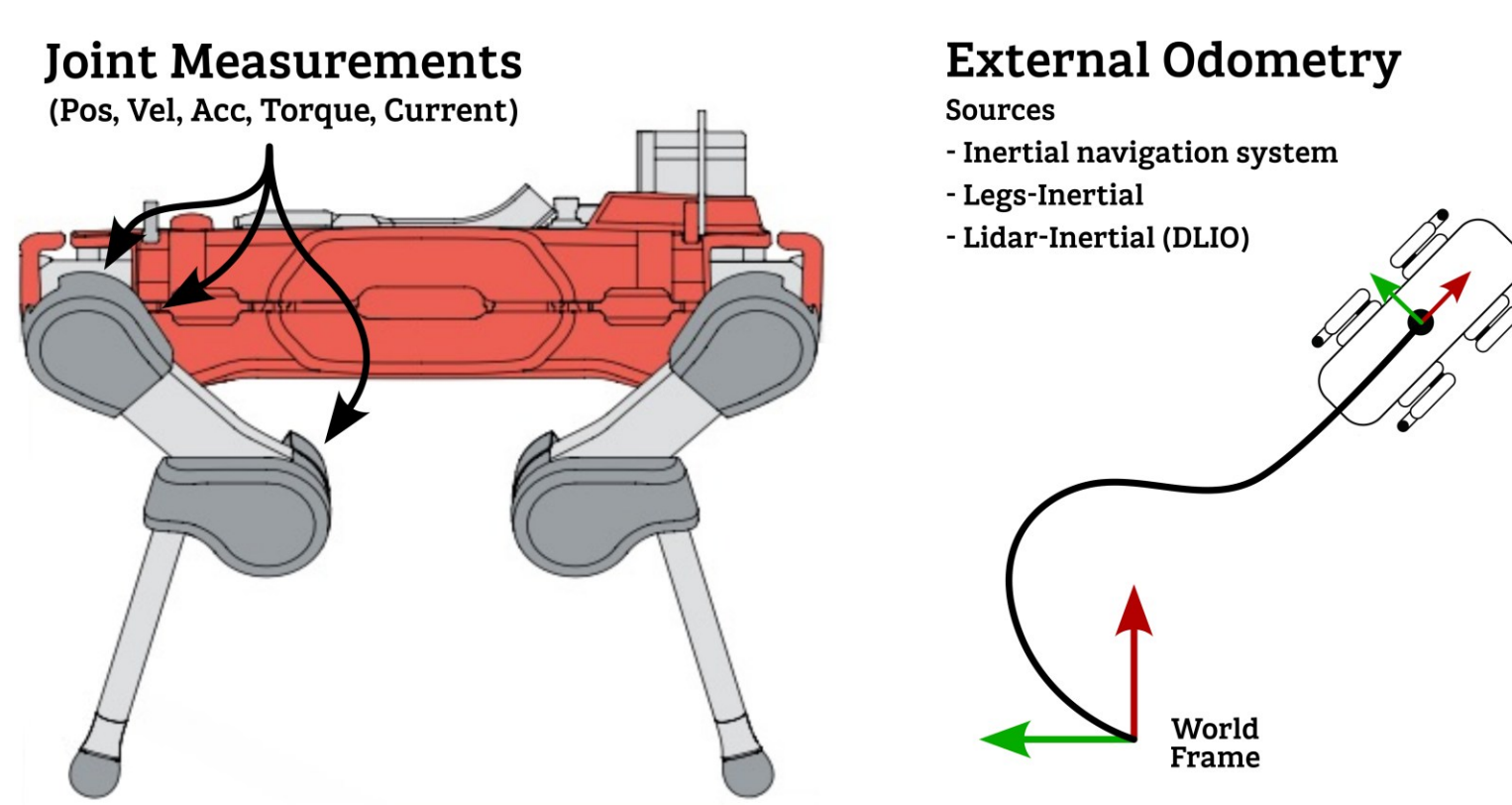
- Quadrupeds walking in soft forest ground must **plan foot placements** to safely avoid obstacles and unstable terrain.
- Traditional **geometric planners assume rigid terrain**, limiting performance on deformable surfaces like mud or vegetation, where contact stability is uncertain..
- We propose **vision-based terrain assessment** to predict contact stability from visual inputs, using **self-supervised** learning from robot experience.
- We derive terrain-dependent stability signals from **proprioceptive** and **odometry** data, providing supervision for training visual predictors.



Terrain Stability from Proprioception

We derive four **physically explainable** terrain-stability metrics based on readily-available proprioceptive data, distilling high-dimensional sensor data into a single **scalar value for every footstep**.

These metrics are specifically formulated to remain **invariant to gait patterns and robot velocity**, isolating the intrinsic physical properties of the environment.



Position Drift

$$\text{avg}(\|\mathbf{p}_t - \mathbf{p}_0\|)$$

Velocity Residual

$$\text{avg}(\|\mathbf{v}_{\text{hip}} + \mathbf{v}_{\text{foot}}\|)$$

Force Consistency

$$\sigma(F_z)/\mu(F_z)$$

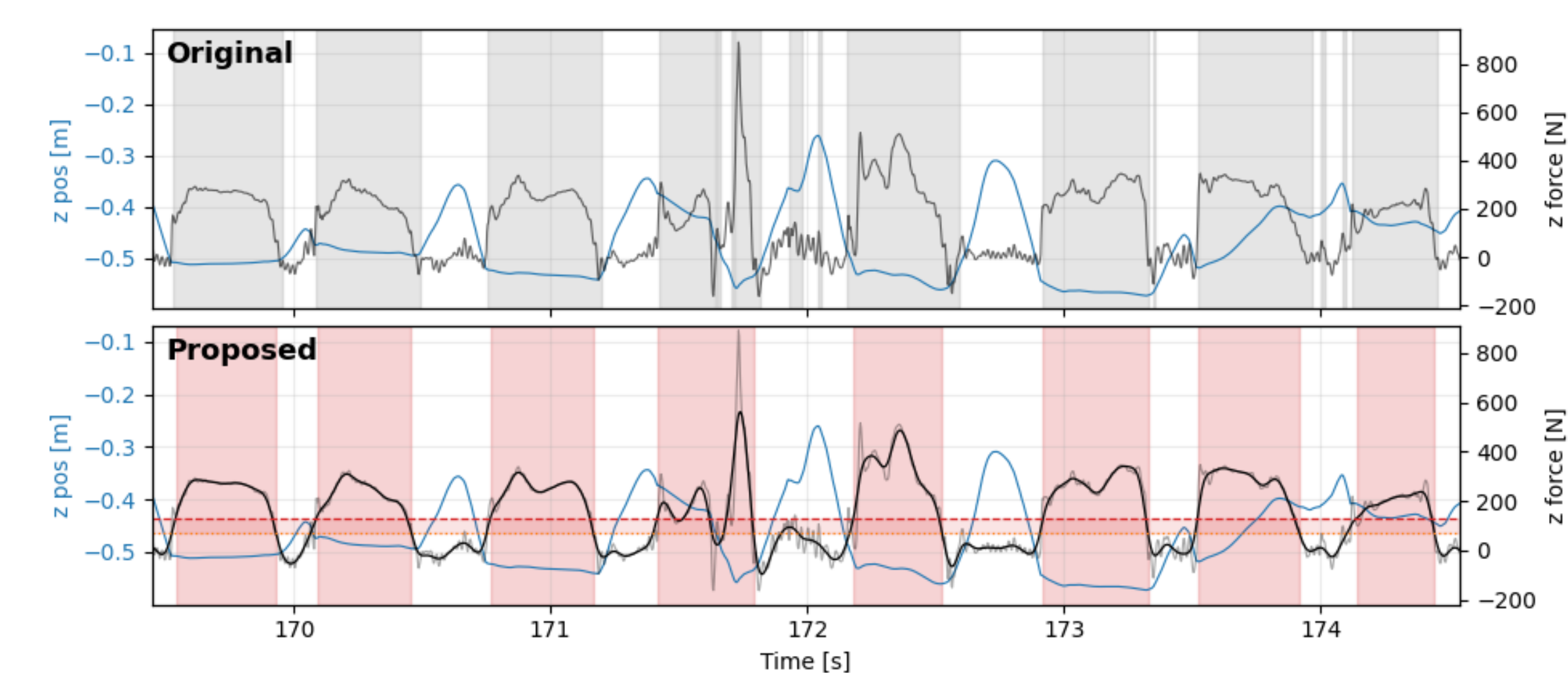
Cost-of-Transport

$$\frac{\sum P_t}{mg \sum \|\mathbf{v}_{\text{hip}}\|}$$

Ground Contact Detection

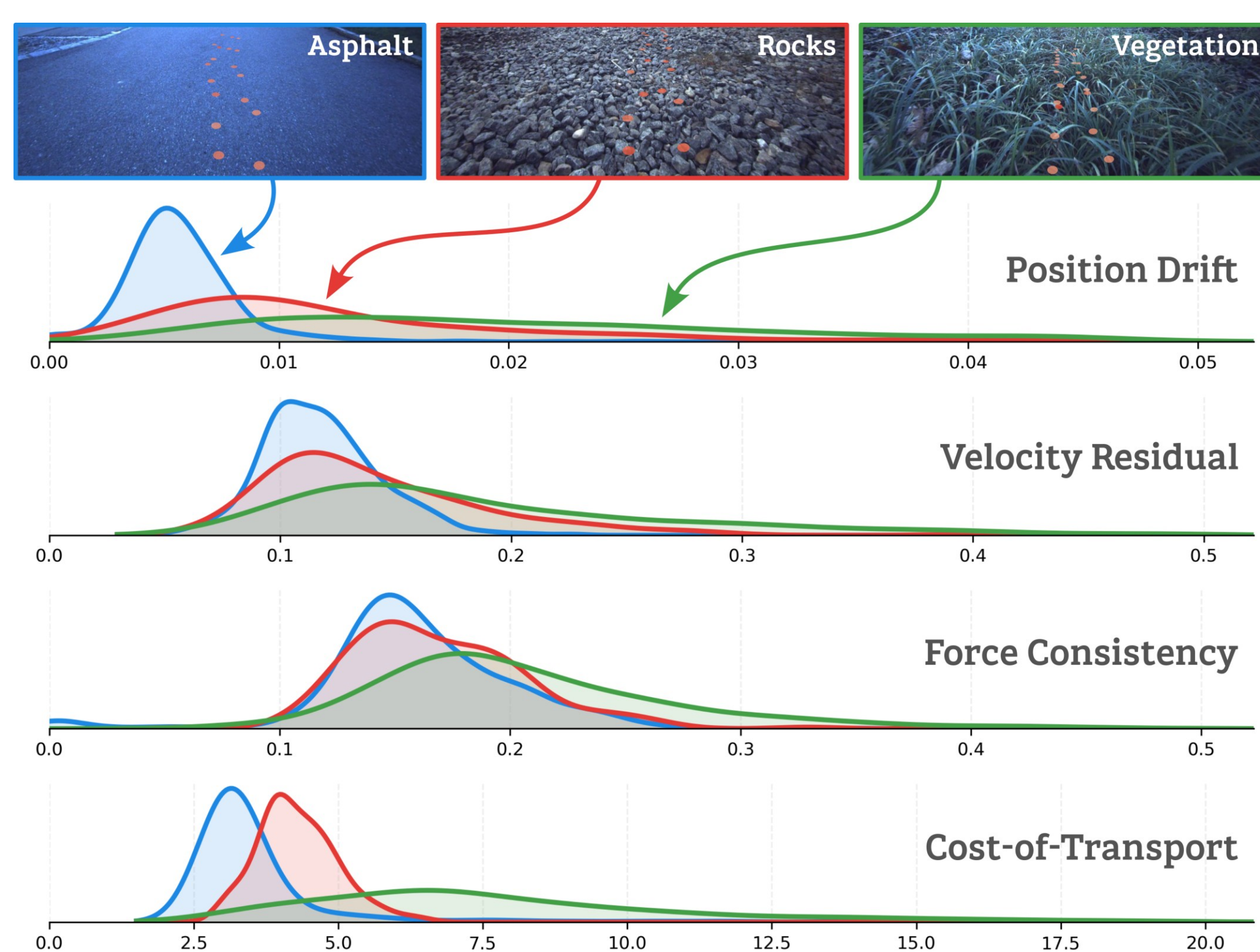
Without feet sensors, detecting contact with ground is not trivial. Embedded contact detector on ANYmal is noisy and unreliable.

We use filtered vertical ground reaction force, passed through a hysteresis threshold and debouncing filter to minimize jitter.



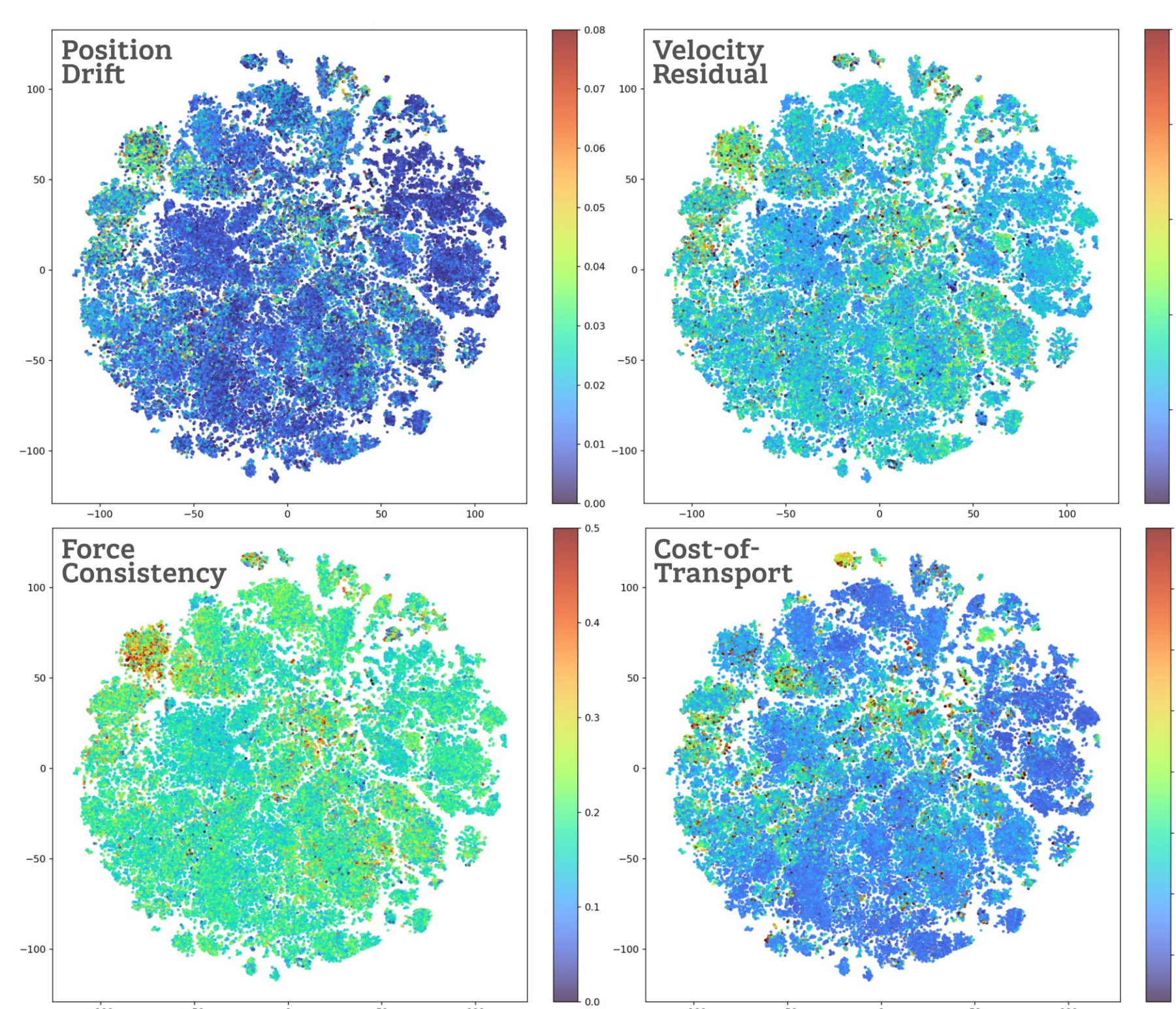
Terrain Differentiation

Analyzing computed metrics on distinct terrains, we notice that **position drift** and **cost-of-transport** can better differentiate between them, based on distributions' below.



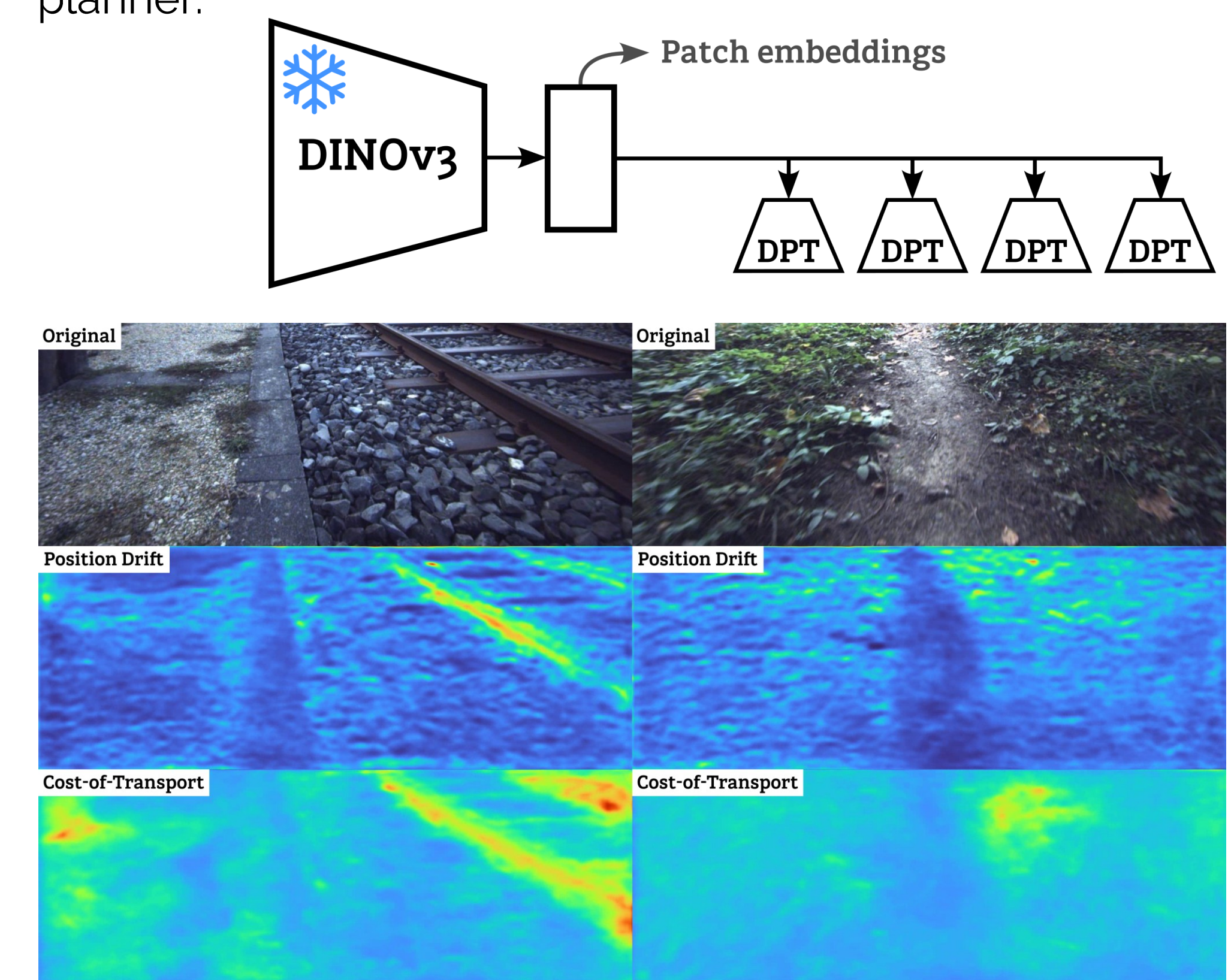
Semantics Correlation

Patch embeddings from DINOv3 [1] to correlate image semantics with our stability metrics. Distinct correlation patterns across the metrics confirm they **capture complementary terrain properties**.



Prediction from Camera

Using a DPT head [2] for dense regression, we train a deep learning based stability predictor that will further be used in closed-loop operation as a low-level footstep planner.



Lessons Learned

- In self-supervised training pipelines, **data curation is key**.
- Lidar-based odometry is **too shaky** for fine-grained drift estimation.
- Inertial navigation system is **unreliable** half the time.
- Proprioception is affected by **gait**, slip on one leg affects others.
- **Entanglement** in dense vegetation is notably harder to detect or predict.

Acknowledgements

This research relies significantly on the GrandTour Dataset [3]. Data used in this work encompasses 24 outdoor runs in diverse environments. We thank the authors for their amazing work.



References

- [1] Siméoni, Oriane, et al. "Dinov3" arXiv preprint arXiv:2508.10104 (2025).
- [2] Ranftl, René, Alexey Bochkovskiy, and Vladlen Koltun. "Vision transformers for dense prediction." Proceedings of the IEEE/CVF international conference on computer vision. 2021.
- [3] Frey, Jonas, et al. "GrandTour: A Legged Robotics Dataset in the Wild for Multi-Modal Perception and State Estimation." arXiv preprint arXiv:2602.18164 (2026).