Preface

To date, the literature offers **no** method that can train a quadruped navigator to move **rapidly** through **narrow or densely cluttered 2-D scenes** while remaining safe. A direct comparison with the latest SOTA—**ABS [1]**, **OCR [2]**, and **Omni-Perception [3]**—underscores this gap: none evaluates (i) multi-room traversal, (ii) dead-end escape, or (iii) tight-corridor passage, all covered in our benchmarks.

ABS illustrates the problem. It succeeds *only* when obstacles are sparse and frontal, showing no exploration or global navigation. Even in moderately wide corridors, ABS exhibits foot–obstacle overlap (Fig. 1), and it simply *refuses to advance* through narrow gaps or half-open doors. Our method removes all of these limitations **efficiently**.

Most of the prior approaches also rely on *prolonged training* or *real-robot fine-tuning*. By contrast, our policy is trained **end-to-end in one hour** and deployed as-is—something the field has never reported.

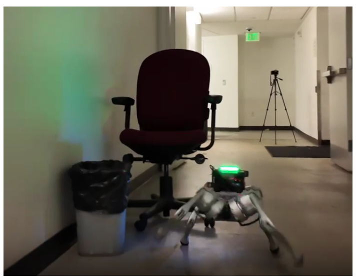

Figure 1: **ABS:** unsafe foot–obstacle overlap in a sparse scene (screenshot from the official video).

Training time itself is not an innovation; it is the *symptom*. That *one hour* suffices to equal or exceed SOTA forces a re-examination of prevailing methodology. No public paper, to our knowledge, solves tasks of comparable difficulty under the same budget (see our "One-take" video).

Contributions

If our success were "just reward functions," *why have those rewards never produced these results?* We identify three pillars:

1. **High-pressure environment design.** High obstacle density, narrow corridors, and long-range exploration appear *together*, enforcing real generalisation.
2. **Safety-driven reward suite.** Exploration reward *up-weights* high-risk experience, accelerating the learning for dense-obstacle avoidance. We enlarge the collision shape and inject observation/action noise for safety, and add a *Goal-Dwell* to encourage both reaching and staying at the goal.
3. **High-efficiency training protocol.** An interaction scheme engineered from first principles yields SOTA-level performance after **one hour** of simulation.

These pillars establish a new baseline: carefully crafted interaction can deliver superior navigation in a fraction of previous training time.

Reviewer 79C1

- **3-D terrain.** Easily added via semantic/geometry cues and a higher-level planner; our module slots into a full 3-D stack unchanged.
- **Other low-level controllers.** Planned (e.g. UAVs); requires retraining with a new environment and I/O.
- **SLAM failures.** Omitted only for brevity: slow footage, dynamic-obstacle misses, mapping errors, and stalls on unseen obstacles.

Reviewer oJwB

- **Reward details.** Will be added; exploration drives toward the *longest ray* or the goal.
- **Safe exploration.** Enlarged shape + noise (safety) and dense- obstacle incentives + Goal-Dwell (exploration).
- **ABS comparison.** ABS itself targets safe navigation; its poor generalisation makes the contrast instructive.
- **Hierarchical control.** Not novel; our innovations lie elsewhere.
- **Frequencies.** 15 Hz LiDAR, 50 Hz low-level inference.
- **Frames.** Commands in robot frame; Fig. 5 global.
- **SEA-Nav w/o Goal-Dwell.** Resets on first goal contact.
- **Action-constraint loss.** Added to PPO and updated in lockstep.
- **SLAM usage.** Breezy-SLAM for odometry only; visual odometry was too drift-prone.

Reviewer HXSn

- **Action constraints.** Auxiliary loss bounding the action space; more effective than clipping/rate limits.
- **Observation design.** Rays alone are not the secret—the *interaction design* is.
- **Encoder necessity.** Handing the full history to the actor blurs which timestep matters. We give the actor only the current observation.
- **Open source.** Code will be released upon acceptance.
- **Why ABS fails.** Training environment too simple; lacks dense layouts and efficient interaction.

Reviewer 5f6v

- **Self-collision.** Pre-trained locomotion policy guarantees none; navigation commands cannot induce it.
- **Point-mass abstraction.** Mitigated by a tracking reward and an action-constraint loss.
- **Ray fidelity.** Finer resolution possible; at $\pi/30$ close objects still intersect at least one beam.
- **Backward motion.** Fully supported; not shown for brevity.
- **ABS performance.** Fails in dense scenes; our videos vs. ABS website illustrate the difference.
- **Full velocity commands.** Enable speed modulation and plug directly into standard velocity-tracking controllers.
- **SLR.** Pre-trained locomotion policy; no joint training needed.
- **2-D from 3-D LiDAR.** Some loss is accepted; 3-D scenes are outside current scope.

References

[1] T. He, C. Zhang, W. Xiao, G. He, C. Liu, and G. Shi, "Agile But Safe: Learning Collision-Free High-Speed Legged Locomotion," in *Proc. Robotics: Science and Systems (RSS)*, 2024.

[2] A. Lin, S. Peng, and S. Bansal, "One Filter to Deploy Them All: Robust Safety for Quadrupedal Navigation in Unknown Environments," *arXiv preprint arXiv:2412.09989*, 2024.

[3] Z. Wang, T. Ma, Y. Jia, X. Yang, J. Zhou, W. Ouyang, Q. Zhang, and J. Liang, "Omni-Perception: Omnidirectional Collision Avoidance for Legged Locomotion in Dynamic Environments," *arXiv preprint arXiv:2505.19214*, 2025.