ICUAS 2026

From Language to Logic (FL2L)

A Theoretical Architecture for VLM-Grounded Safe Navigation

Kristy Sakano | Kalonji Harrington | Matthew Bandos | Jianyu An | Mumu Xu

COLLEGE OF COMPUTER, MATHEMATICAL, & NATURAL SCIENCES

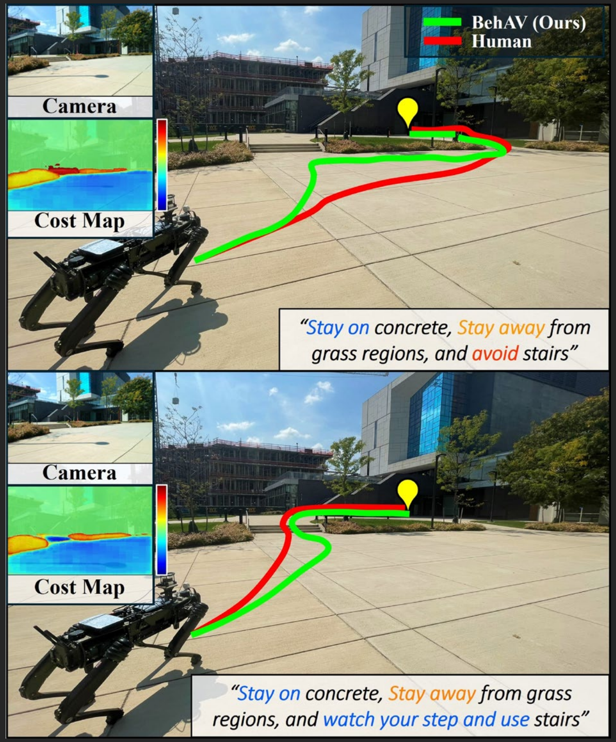
A. JAMES CLARK SCHOOL OF ENGINEERING

2026 INT'L CONFERENCE ON UNMANNED AIRCRAFT SYSTEMS
ICUAS 2026
uasconferences.com

Vision + Mobile Robots = Robots That Can See AND Navigate!

Current technology supports vision-guided navigation of fixed-wing aircrafts, quadcopters, and more.

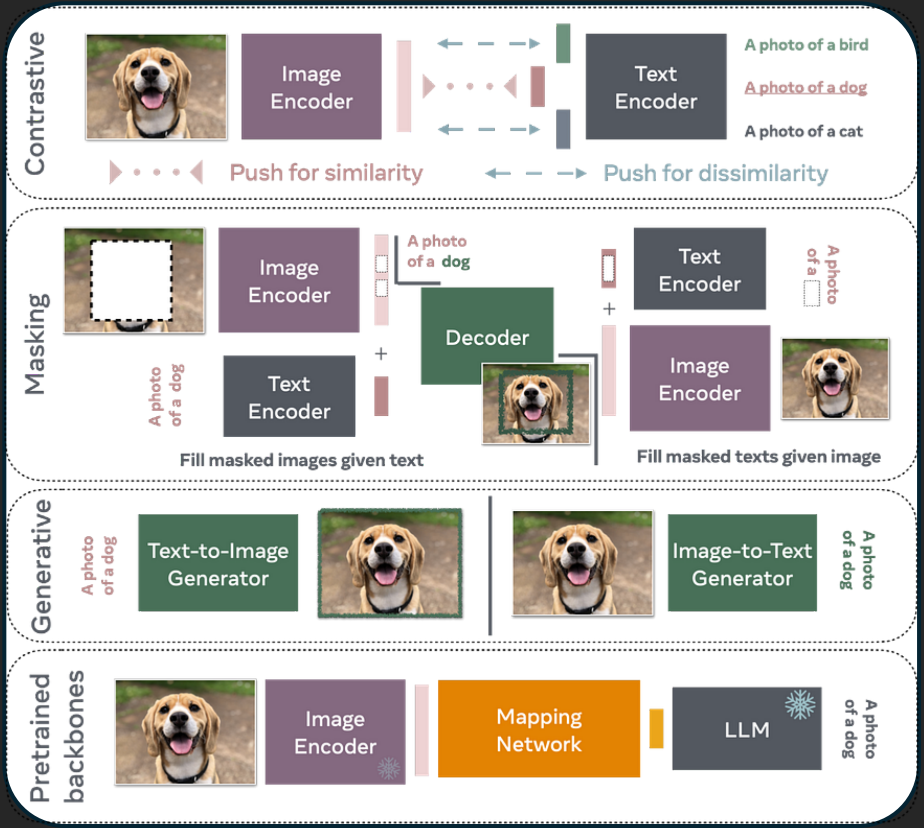
Song et al. (2024)



Weerakoon et al. (2024)

Vision-Language Models (VLMs)

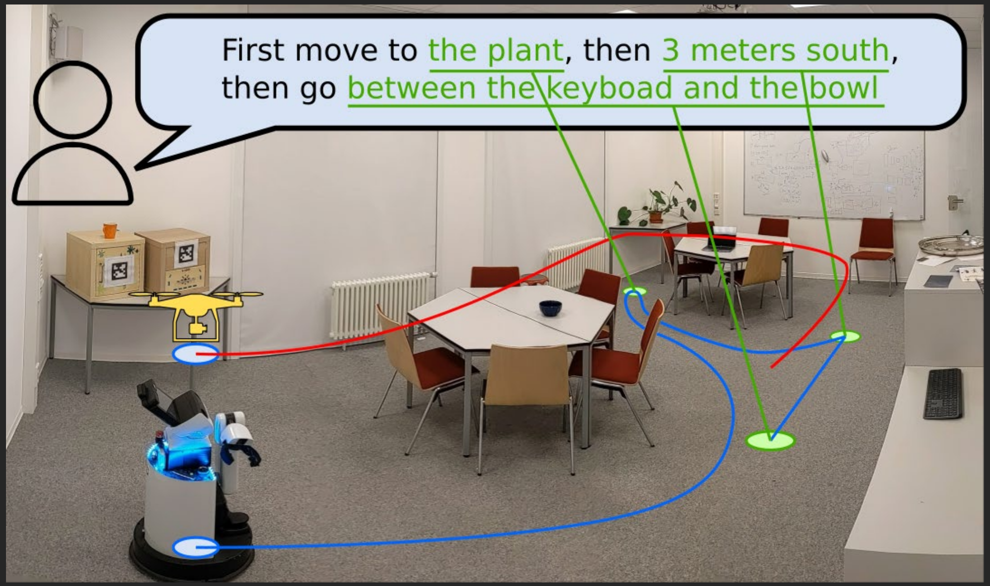
Multi-modal AI that can process, understand, and reason about visual data (image and videos) and text.



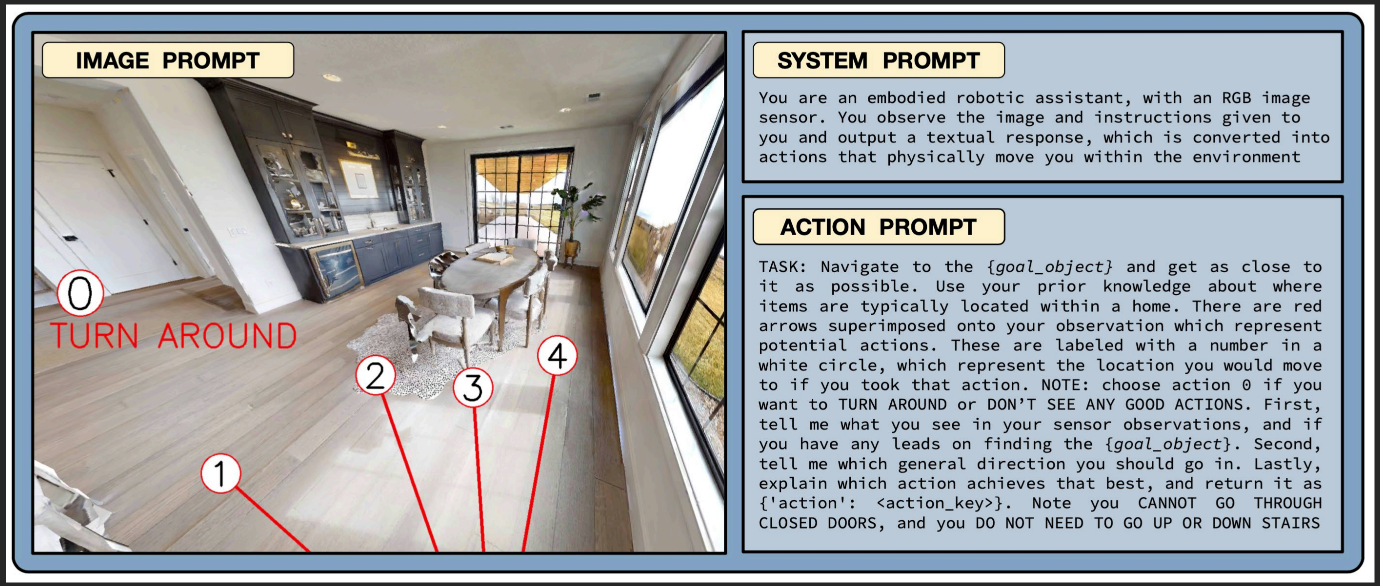
Bordes et al. (2024)

VLMs for Mobile Robot Navigation

Navigating an environment with a RGB camera, while parsing instructional text from a human.



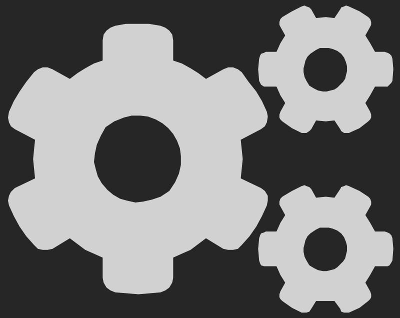
Huang et al. (2023)



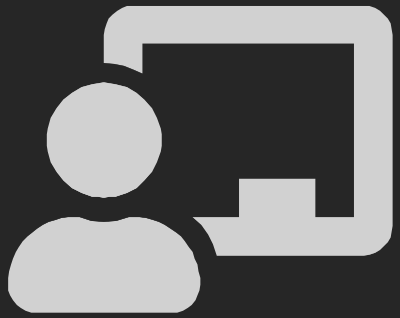
Goetting et al. (2024)

Verification and Validation (V&V) of Mobile Robots

How do we ensure that autonomous mobile robots are safe and following the rules?



Empirical Testing
demonstrations
scenarios
high-fidelity simulations



Safety Heuristics
time-to-collision
near-misses
minimum separating distance

Formal Logic & Rules
finite state machines
hard-coded rules
formal specifications

Bridging Human Intent with Mathematical Guarantees

FL2L blends the adaptability of VLM-based navigation with the mathematical rigor of Formal Methods.

Mission Operator

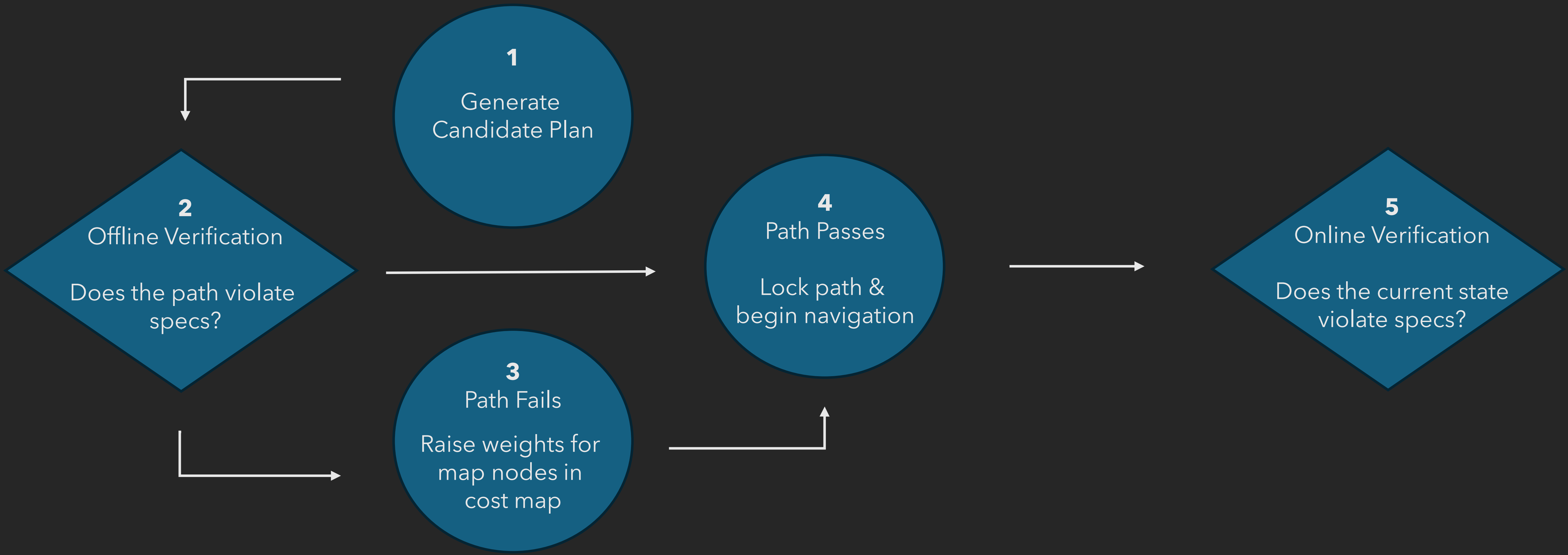
"If you go into the grass area, then eventually (within 3 secs), you must come out of the grass and stay above the sidewalk for at least 5 secs."

Offline Verification

Online Verification

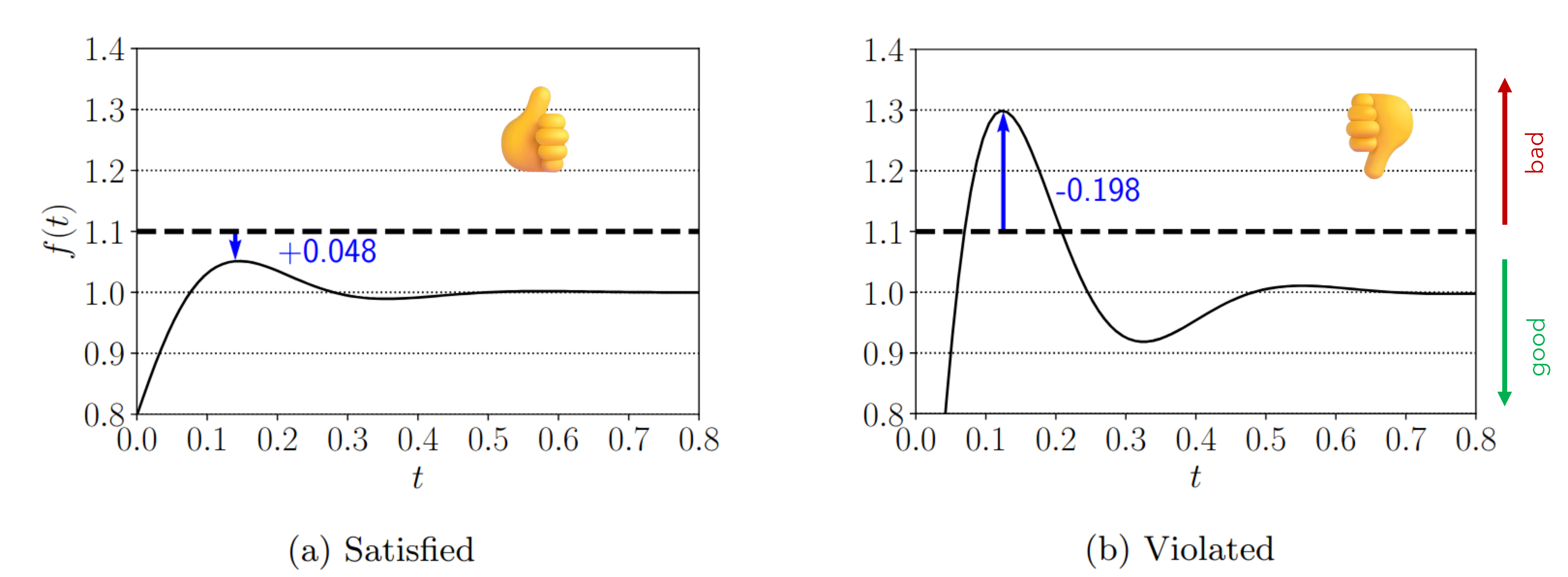
The Formal Evaluation Loop

The system maintains a mathematically proven, safe path to the goal.



Quantifying Safety with Signal Temporal Logic

The core concept: moving beyond safety as a "true" and "false" concept for dense-time signal.



Yamaguchi, Hoxha, and Nickovic (2025)

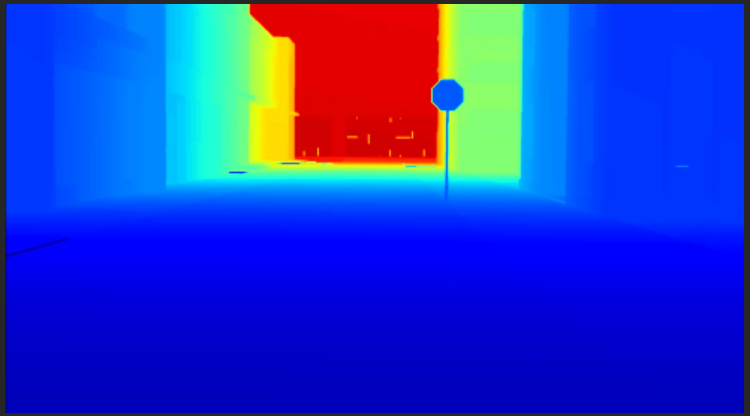
Example rule: *"Always stay below 1.1"* → G[0,∞] f(t) ≤ 1.1

Virtual Proving Ground: Verifying Our Approach in Simulation

Combining Blender (environment modeling), IsaacSim (robot sim), and CLIPSeg (VLM)

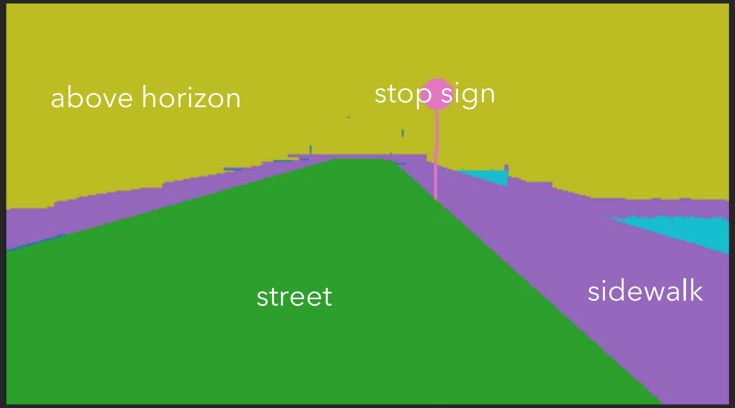
RGB Image
(simulated in Blender)

→



Monocular Depth Estimation

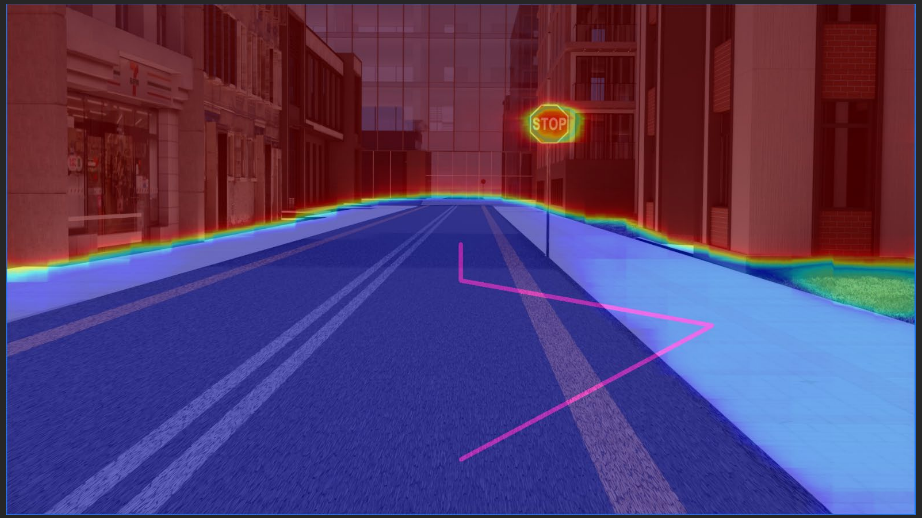
+



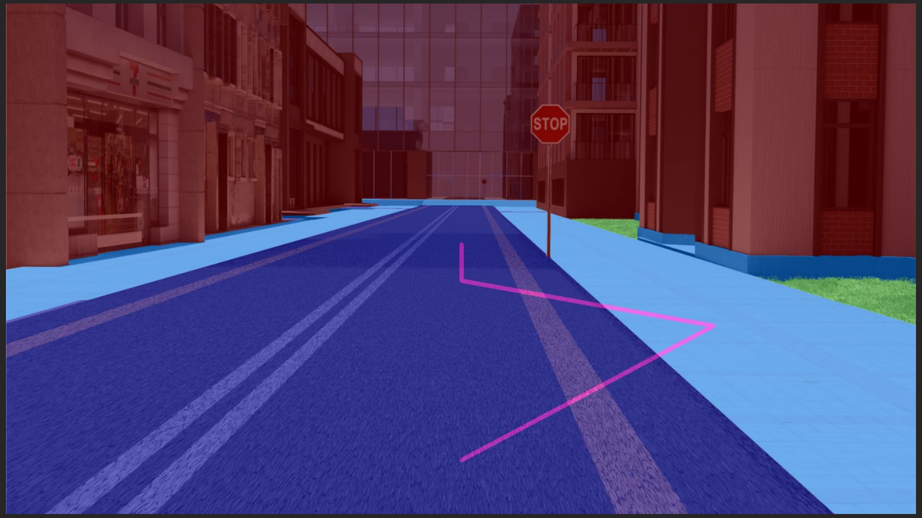
Object Segmentation

Virtual Proving Ground: Verifying Our Approach in Simulation

Combining Blender (environment modeling), IsaacSim (robot sim), and CLIPSeg (VLM)

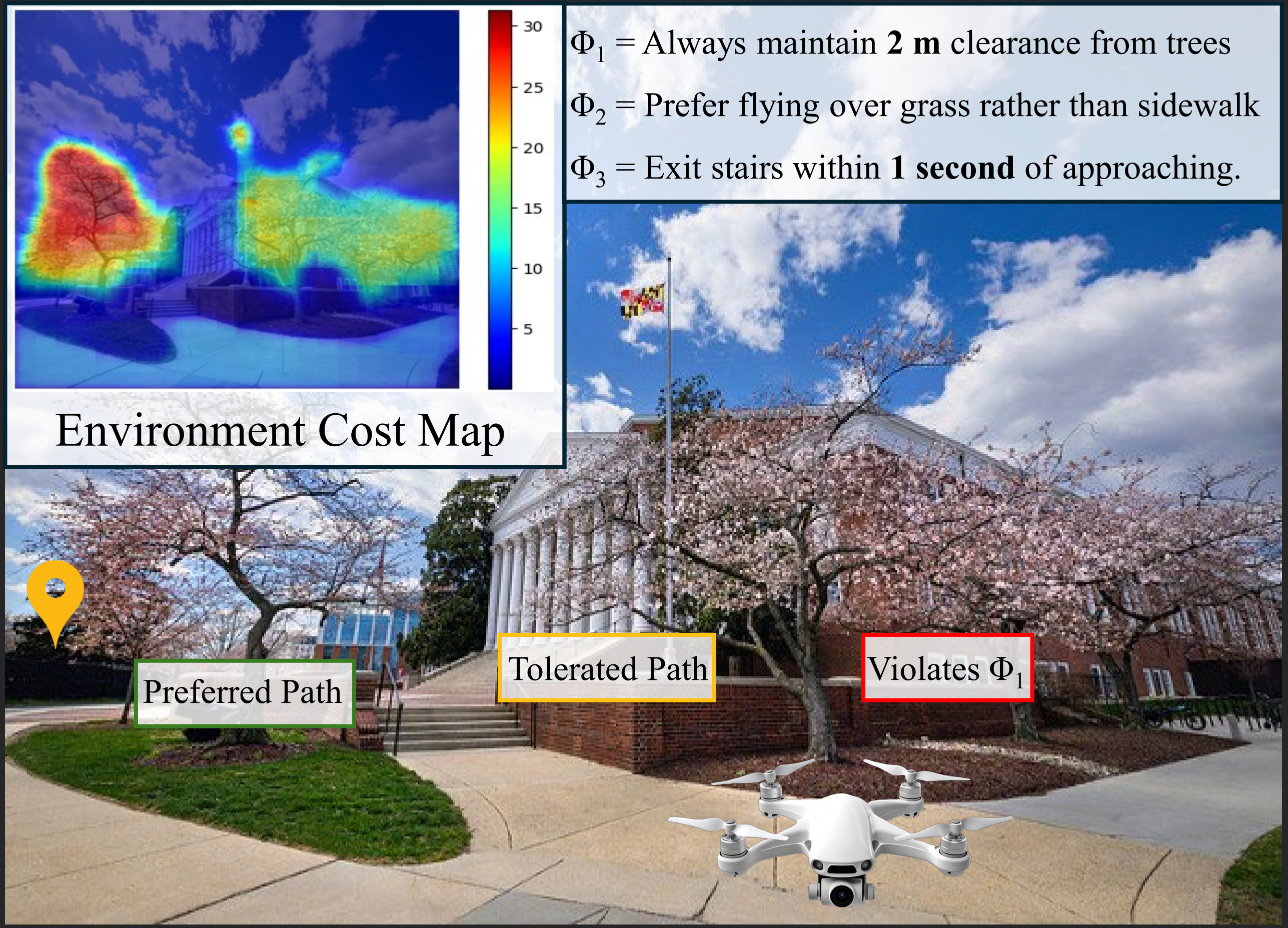


Predicted Cost Map
(CLIPSeg)



Ground Truth Cost Map

From Language to Logic (FL2L): A Theoretical Architecture for VLM-Grounded Safe Navigation



The Team

Dr. Huan Xu
PI

Kristy Sakano
PhD Student

Kalonji Harrington
MS Student

Jianyu An
PhD Student

Matthew Bandos
Undergraduate Student

References in this Talk

Links to references used in this talk.

Slide 2 “Vision + Mobile Robots”: (a) https://arxiv.org/abs/2404.00210 (b) https://arxiv.org/abs/2409.16484
Slide 3 “Vision-Language Models (VLMs)”: https://arxiv.org/abs/2405.17247
Slide 4 “VLMs for Mobile Robot Navigation: (a) https://vlmaps.github.io/ (b) https://jirl-upenn.github.io/VLMnav/
Slide 14 RTAMT https://github.com/nickovic/rtamt

Font Awesome