

PROJECT PORTFOLIO · 2024-2026

Gergely Dinya

Computer vision engineer working on 3D reconstruction, multi-camera tracking and instance segmentation, with published research papers and open-source repositories.

MSc in Computer Science (AI specialization)

github.com/gergelydinya · gergelydinya.com

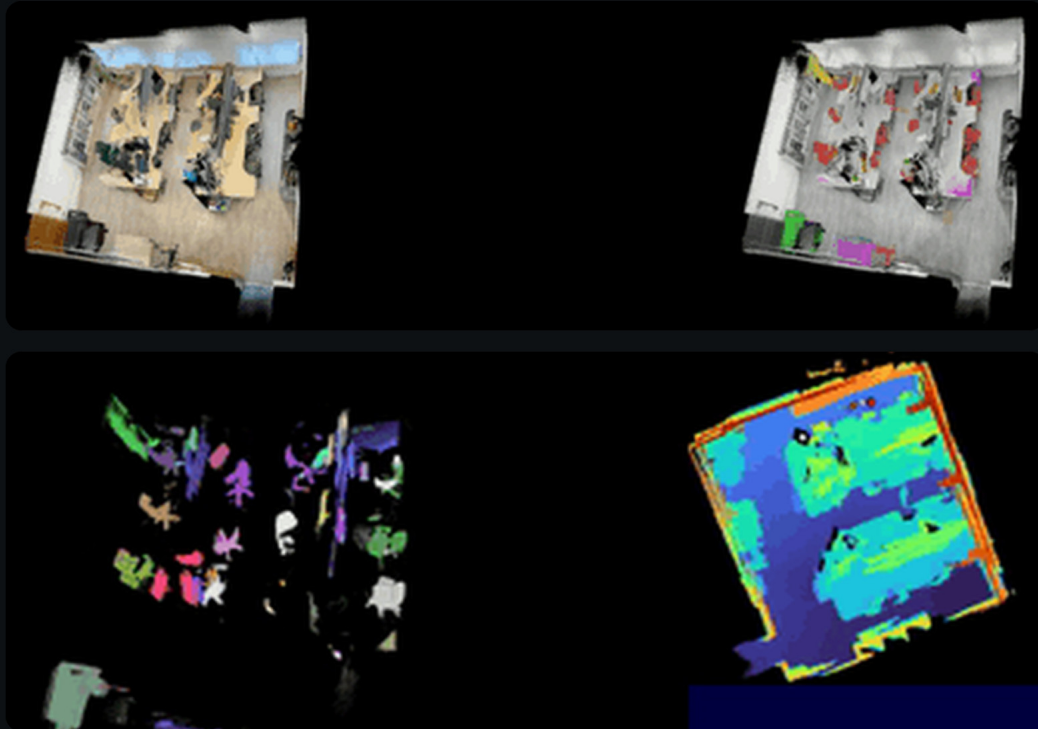
linkedin.com/in/gergely-dinya

dinyagergely@gmail.com

Available from November 2026 · EU citizen



SceneVGGT

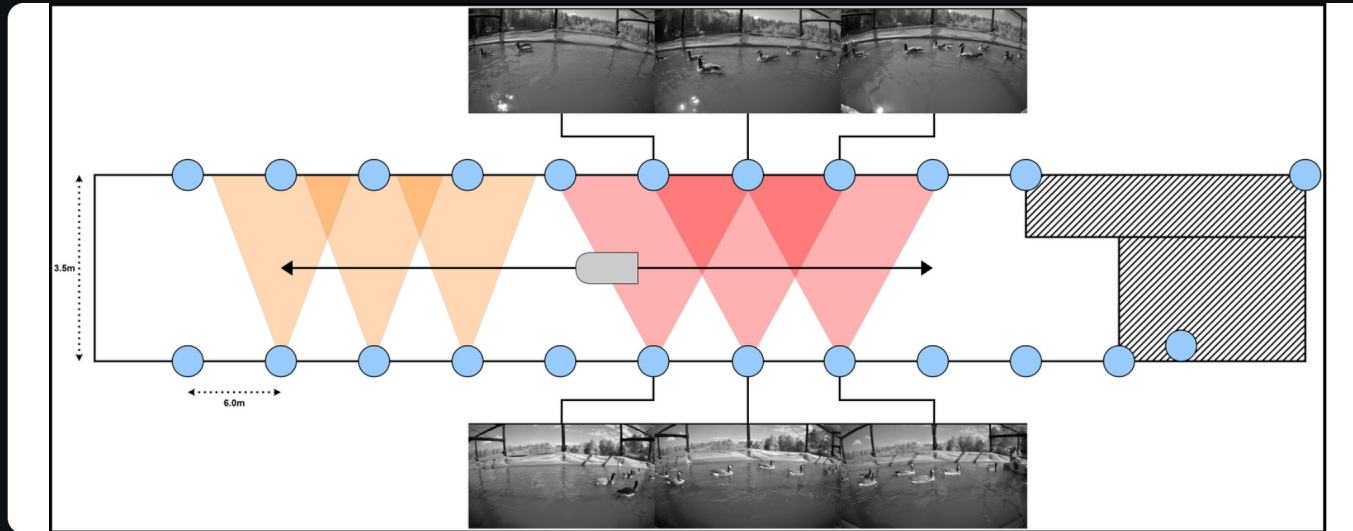


Top: RGB reconstruction and closed-set semantic segmentation.
Bottom: 3D instance memory and navigation map.

- An online, end-to-end semantic SLAM system adding LiDAR-grounding to a VGGT backbone.
- The system processes the data stream in a sliding-window fashion, creating and aligning submaps to keep VRAM usage constant.
- 2D closed-set semantic instance masks are lifted to 3D, tracked and stored in a temporal memory for change detection.

Co-first author · [paper](#) · [code](#)

AviTrack



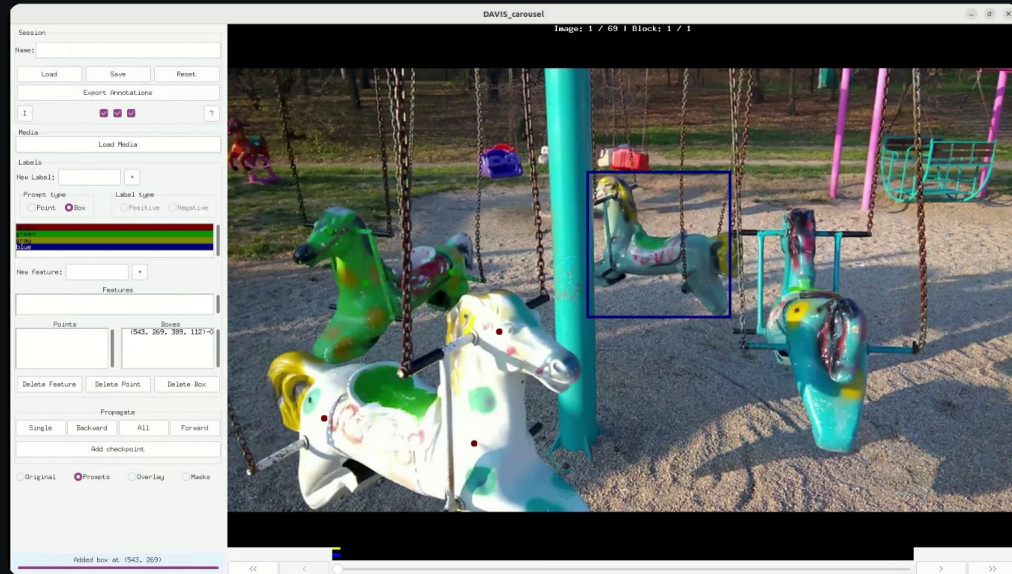
Aviary layout: camera pairs mounted on opposite sides of a floating structure, with 12–16 geese swimming between them. The goal was to ensure that Goose A in one camera remained Goose A in all cameras.

Co-first author · [paper](#) · [code](#)

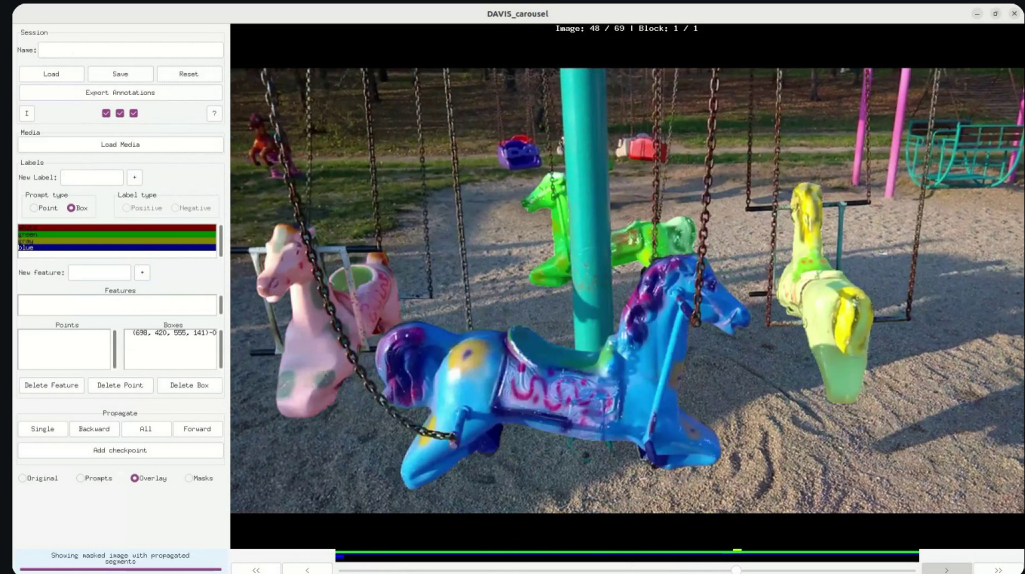
Gergely Dinya · Project portfolio

- 01 **Orientation estimation:** as the cameras are mounted on a flexible structure, their roll and pitch need to be estimated at the beginning of each session.
- 02 **Synchronise:** refine the second-level camera synchronisation to avoid drift in detections, and thus in position estimates, across cameras.
- 03 **Detect:** fine-tuned YOLO models detect the animals, and their collars where visible.
- 04 **Cross-view instance association:** using visual and positional cues.
- 05 **Track:** SORT-based position tracking for initial sub-tracklet generation.
- 06 **Final pairing:** sub-tracklets are merged into final tracks using a combined-metric calculation and linear assignment.

SAMannot



1 · Users prompt the first frame of the sequence.



2 · Segmentation masks are propagated with SAM2, and new prompts are calculated automatically for the next image sequence.

An efficient video annotation tool built around Segment Anything 2, designed to reduce the human attention needed during annotation as much as possible.

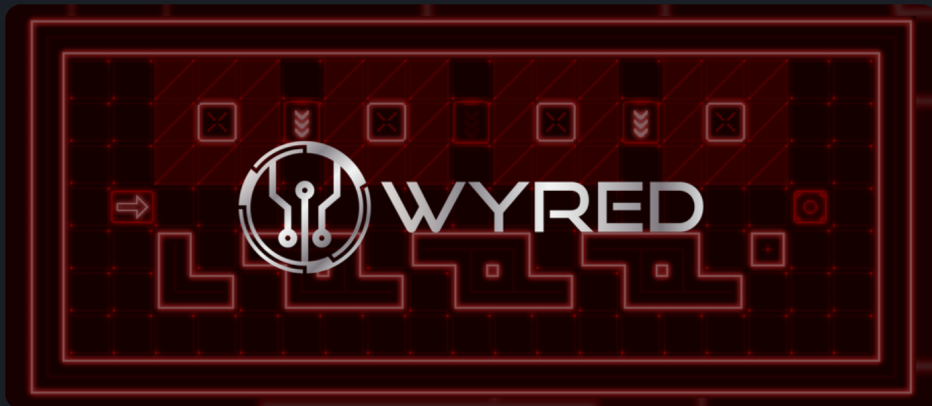
First author · [paper](#) · [code](#)

Gergely Dinya · [Project portfolio](#)

Wyred and the Titan framework

Wyred

A 2D puzzle game, released on Steam, about routing wires to form a circuit. The main hook is that players cannot cross previously laid tracks, so they have to plan routes ahead and adapt as they go.



[Steam page](#)

Titan

A lightweight C++ game development framework with OpenGL rendering. Its subsystems - aimed at helping with resource management, shader handling, audio, etc. - are either written from scratch or wrap existing libraries.

USED IN

[Wyred](#), released

[Pixel Armada](#), under development

[code](#)