PhD projects (Computer vision)

- Urban AI Shields: GeoAI and Multimodal Data for Predicting Extreme Weather Health Impacts on Vulnerable Urban Populations
- Generative World Models for Unsupervised Procedural Graph Discovery

Urban AI Shields: GeoAI and Multimodal Data for Predicting Extreme Weather Health Impacts on Vulnerable Urban Populations

Supervisor: Yijing Li

Project Description

Cities are on the frontline of the climate crisis. Heatwaves are intensifying, extreme rainfall events are growing more frequent, and the populations least able to adapt: the elderly, low-income communities, people with chronic illness, and those with limited social support, bear a disproportionate burden. This project will develop a next-generation GeoAI framework to understand and predict how extreme urban weather events such as heatwaves and rainstorms impact the health and resilience of vulnerable populations, and to identify actionable pathways to reduce healthcare demand before crises unfold.

You will work at the intersection of artificial intelligence, urban geography, public health, and climate science, harnessing the power of multi-source, multimodal data to generate insights that are spatially precise, socially grounded, and operationally useful. The project is expected to build up a dynamic vulnerability mapping platform integrating real-time remote sensing and AI inference, to facilitate the integration of model predications on healthcare demand surges under extreme weather scenarios, aiming to provide evidence-based recommendations for urban planners and public health authorities. We are looking for candidates who are strong in quantitative or computational skills (Python, R, deep learning frameworks), interest in spatial data, remote sensing, or urban analytics, but also with passion for climate justice and health equity, and curiosity that crosses disciplinary boundaries.

Generative World Models for Unsupervised Procedural Graph Discovery

Supervisor: Luis C. Garcia Peraza Herrera

Project Description

This computer vision PhD project focuses on neuro-symbolic video understanding: transitioning from dense spatio-temporal perception to extracting the stochastic, branching topologies embedded in procedural video. Current self-supervised video foundation models optimise continuous latent manifolds that fundamentally fail to capture the discrete, hierarchical logic dictating complex workflows.

This project bridges the critical gap between continuous feature learning and discrete algorithmic cognition. Building on our prior success in self-supervised learning for procedural video [1], the objective is to design unsupervised frameworks that autonomously discover executable probabilistic finite state machines directly from large-scale, unannotated video archives. By integrating vector-quantised bottlenecks and unsupervised graph induction with temporally aware representations, the framework will discretise continuous visual embeddings into interpretable directed acyclic graphs. This extracted logic will subsequently condition a generative video world model, enabling the AI to simulate counterfactual procedural branches and forecast spatio-temporal states zero-shot.

While advancing foundational representation learning, these architectures strictly target high-stakes domains (e.g., surgical automation). By eliminating the annotation bottleneck of supervised phase recognition, this research yields verifiable state transitions for AI decision-making. Extracting these underlying procedural topologies will serve as the cognitive engine for generative planning, providing the verifiable safety architecture required for autonomous robotics.

References

[1] C. Che, C. Wang, X. Chen, S. Tsoka, and L. C. Garcia-Peraza-Herrera, "A Stitch in Time: Learning Procedural Workflow via Self-Supervised Plackett-Luce Ranking," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2026, pp. 17000-17010.