Limit-Aware Hybrid AI for Reliable Deep Perception

The deployment of autonomous systems in safety-critical environments faces a persistent reliability challenge. While deep learning models have achieved high performance on standard benchmarks, they frequently fail to generalize to the physical and regulatory constraints of the real world. Data-driven models often produce predictions that violate basic kinematic laws (e.g., impossible acceleration) or geometric consistency (e.g., object permanence failures). Conversely, current hybrid approaches often incorporate these constraints only as soft loss terms or post-processing filters, which do not guarantee compliance during inference [1], [2].

My research goal is to address this fundamental gap by formalizing Limit-Aware Spatial Intelligence. Rather than treating physical and regulatory limits as external guardrails, I propose a framework where these constraints are embedded as structural components of the perception architecture itself. By encoding mathematical limits directly into the attention mechanisms and representation layers of vision models, I aim to develop systems that are not only statistically accurate but operationally robust and interpretable.

Background: Foundational Insights from Transportation Safety

My doctoral research addressed transportation safety challenges through two distinct but complementary tracks: data-driven behavioral modeling and geometry-informed computer vision. This dual focus provided me with first-hand experience in the promises and pitfalls of each approach, forming the empirical basis for my future research.

Data-Driven Modeling of Driver Behavior:

In **Padmanaban et al.** [3], I introduced a novel speed-adjusted jerk thresholding method to engineer features that are grounded in vehicle kinematics and driver operational characteristics. Through a rigorous validation of the implemented machine learning pipeline, the best model identified aggressive driving with 94% accuracy, significantly outperforming standard baselines. I extended this in **Padmanaban et al.** [4] by performing a comparative analysis of acceleration and deceleration profiles, contrasting real-world aggressive driving with EPA fuel economy test cycles to reveal significant behavioral and environmental implications. This work demonstrated that even powerful data-driven models are most effective when their feature space is constrained by domain-specific physical principles.

Geometry-Informed Computer Vision for Cycling Safety:

In **Padmanaban et al.** [5], I developed a geometry-informed computer vision system for detecting vehicle overtaking events from bicycle-mounted cameras, achieving 98.7% precision by leveraging perspective geometric principles rather than relying on data-driven appearance models. Building on this, my work in **Padmanaban et al.** [6] established a geometry-based passing distance estimation method that achieves a mean absolute error of just 9.5 cm using only monocular vision features. My current and final phase of doctoral work integrates these components to investigate how fine-grained vehicle types (e.g., SUVs vs. sedans) systematically influence overtaking behavior. This research proved that explicit mathematical models of geometry can provide robust and computationally efficient solutions for safety-critical estimation tasks.

Key Insight: My research experience established that while data-driven models excel at feature extraction and classification, they lack inherent physical reasoning. Conversely, physics-based methods are precise but can be unreliable if not integrated with robust learning. As a next step, I intend to develop an independent research program that will synthesize these two streams: using the architectural flexibility of deep learning to learn from data, while structurally enforcing the physical and geometric limits.

Future Research Agenda: Limit-Aware Hybrid AI

My long-term vision is to transform perception from a pattern-matching task into a trustworthy, rigorous engineering discipline where AI systems explicitly quantify their own reliability. I propose a unified Limit-Aware Perception Framework illustrated in **Figure 1**, that interleaves three distinct pathways of reasoning(data-driven, physics-guided, and knowledge-guided), to create systems that are "aware" of their operational boundaries.

Thrust 1: Physically Grounded Architectures (The "How")

Purely data-driven models can hallucinate physically impossible trajectories because they lack structural constraints.

Proposed Approach: I will develop architectures that embed vehicle dynamics directly into the model's training and inference. By integrating fundamental kinematic constraints (e.g., friction-limited acceleration $|a| \leq 8 - 10\ m/s^2$, emergency braking limits $\approx 9\ m/s^2$, and non-penetration geometry) directly into the attention mechanism, my framework acts as a differentiable constraint mechanism. Unlike approaches relying on soft loss penalties, these constraints are implemented as hard attention masks (M_limit) that structurally compel the model's representations to remain consistent with physical laws [7]. This ensures that even in the presence of sensor noise or occlusion, the model's hypothesis space is bounded by physical reality.

Thrust 2: Geometry-Informed Representation Learning (The "Where")

Current perception models often struggle with spatial reasoning from 2D inputs, relying on massive datasets to approximate depth.

Proposed Approach: Building on my doctoral success in calibration-free estimation, I will research methods to tokenize geometric invariants, such as horizon lines, camera intrinsics, and vanishing points, and fuse them with visual features in the latent space. This approach explicitly grounds the visual representation in projective geometry, enabling monocular systems to estimate distance and velocity with the fidelity required for safety applications (e.g., precisely tracking a cyclist's lateral distance) but with the computational efficiency of standard 2D networks.

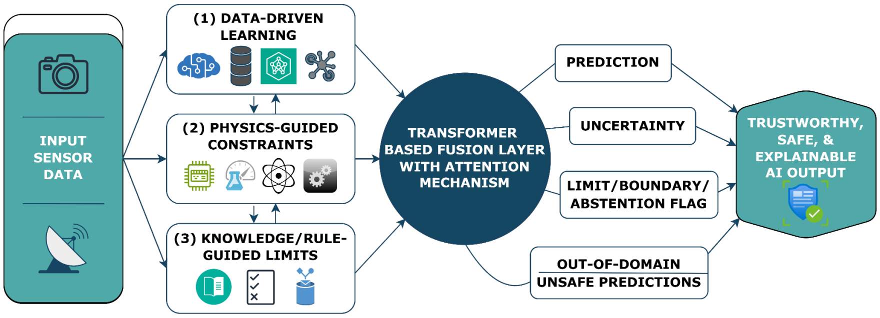

Figure 1. Limit-Aware Hybrid AI framework embedding physics-guided constraints across feature, transformer, and output layers.

Thrust 3: Operational Awareness and Structured Abstention (The "When")

A truly trustworthy system must know when not to act. Current "black box" models cannot distinguish low-confidence noise from physically impossible predictions.

Proposed Approach: I will introduce *Boundary Tokens*, which are learnable embeddings that encode explicit limits such as maximum legal speeds, sensor effective ranges, or lane boundaries. These vectors act as "sentinels" in the attention mechanism, allowing the model to explicitly reason about violations. Instead of a single scalar confidence score, the system will produce structured abstention signals: explicit, interpretable flags indicating why a prediction is unreliable. This transparency allows downstream planning systems to make informed, safety-critical decisions, such as triggering safe mitigation strategies or requesting human intervention when safety guarantees cannot be met.

Broader Impact: Advancing the Science of Perception

My research establishes the foundations for Visually Grounded Spatial Intelligence, a new class of perception systems that possess an innate understanding of the 3D physical world. By transforming perception from a statistical pattern-matching task into a physically constrained reasoning process, this work addresses the most critical bottleneck in modern computer vision: the gap between high-accuracy object detection and actionable, safety-critical understanding.

Transforming Perception Systems: The immediate impact of this work will be the creation of perception frameworks that are not only accurate but certifiable. By embedding hard physical and geometric limits into the architecture, we move closer to perception systems that can guarantee consistency by ensuring, for example, that a tracked object never violates the laws of physics. This reliability is the prerequisite for moving autonomous agents from controlled pilots to widespread, unsupervised deployment in complex urban environments.

Democratizing Spatial Intelligence: Beyond transportation, this "Limit-Aware" paradigm offers a scalable blueprint for any domain where computer vision intersects with the physical world. Whether for industrial robots collaborating with humans or smart infrastructure monitoring vulnerable road users, the need for perception that understands its own boundaries is universal. To accelerate this shift, I am committed to Open Science and will release the "Limit-Aware Perception Library." This open-source toolkit will allow the broader engineering community to integrate physics-based constraints into standard detection models (like YOLO and DETR), lowering the barrier to entry for building trustworthy, spatially intelligent systems.

Key References

[1] F. Zhu, X.-Y. Zhang, Z. Cheng, and C.-L. Liu, "Revisiting Confidence Estimation: Towards Reliable Failure Prediction," arXiv preprint, 2024.

[2] F. Zhu, Z. Cheng, X.-Y. Zhang, and C.-L. Liu, "Rethinking Confidence Calibration for Failure Prediction," in ECCV, 2022.

[3] **G. Padmanaban**, F. Feng, E. Dai, S. Ankit, G. Hu, and Y. Zhao, "A Machine Learning Pipeline Framework to Identify Aggressive Driving Based on Vehicle Kinematics and Driver's Pedal Operations," [Under Review], 2025.

[4] **G. Padmanaban**, F. Feng, E. Dai, S. Ankit, G. Hu, and Y. Zhao, "A Comparative Analysis of Acceleration and Deceleration Profiles for Aggressive Driving Styles and Fuel Economy Test Cycles," SAE WCX, 2025.

[5] **G. Padmanaban**, R. Moustafa, and F. Feng, "A Geometry-Informed Computer Vision Method for Detecting and Examining Overtaking Vehicles From A Bicycle," TRB Annual Meeting [Accepted], 2026.

[6] **G. Padmanaban**, and F. Feng, "Geometry-Informed Distance Estimation from 2D Bounding Boxes for Vehicle Overtaking Analysis," [Under Review], 2025.

[7] Z. Liu, Z. Guo, Z. Cen, H. Zhang, J. Tan, B. Li, and D. Ding, "Constrained Decision Transformer for Offline Safe Reinforcement Learning," ICML, 2023.