David Fernandez

dfr522@gmail.com | +1-864-722-7039 | Clemson, SC

Summary

I am a PhD candidate in Computer Science at Clemson University, bridging explainable AI, multimodal machine learning, and security systems for safety-critical applications. My research spans Large Language Models (LLMs), Vision Language Models (VLMs), and deep learning architectures, with four first-authored publications advancing component-level explainability, zero-shot reasoning, and adversarial scenario analysis. As a member of Clemson's VIPR-GS Research Program, I develop hierarchical LLM reasoning frameworks and VLM evaluation systems for the U.S. Army's NGCV program Army's Next Generation Combat Vehicle (NGCV) program. At BMW, I research and develop AI production security frameworks and edge deployment systems. I focus on creating robust, interpretable AI that translates cutting-edge research into real-world impact.

Education

2024 - present	Ph.D. in Computer Science at **Clemson University**	(GPA: 4.0/4.0)
2022 - 2024	M.S. in Data Science and Informatics at **Clemson University**	(GPA: 4.0/4.0)
2010 - 2018	B.S. in Computer Science at **ITAM**	(Mexico City, Mexico)
2010 - 2018	B.S. in Business Administration at **ITAM**	(Mexico City, Mexico)

Work Experience

Virtual Prototyping of Ground Systems (VIPR-GS) for U.S. Army Next Generation Combat Vehicle (NGCV) program, Clemson University 2025 - Present

- Designed novel semantic embedding framework to bridge unstructured visual scenes with structured regulatory specifications, enabling interpretable AI diagnostics
- Developed hierarchical LLM reasoning architecture leveraging prompting strategies and confidence-gated propagation to attribute system failures to specific components, demonstrating that structured reasoning outperforms traditional deep learning approaches in safety-critical anomaly detection
- Integrated VLMs with high-fidelity simulations to evaluate multimodal decision-making, achieving 68.3% agreement with expert decisions versus 61.3% for traditional autopilot systems
- Created MetaVLM evaluation system for vision-language model metadata optimization in driving decision-making contexts revealing strategic metadata selection captures 99.6% of full-information performance while reducing computational overhead by 43% and inference latency by 9%
- Integrated multimodal VLMs with semantic embedding pipelines to enable zero-shot reasoning on unstructured visual inputs aligned with structured knowledge bases, demonstrating 7% improvement over traditional deep learning baselines

AI Researcher at BMW Group 2024 - Present

- Co-authored SACAIR 2025 publication on edge AI deployment strategies, implementing small language model framework with embedding-based selection that enables distributed inference across constrained devices while maintaining production-grade performance
- Designed multilingual semantic document framework using transformer embeddings to automate compliance verification, detecting structural and semantic changes across regulatory documents
- Conducted systematic security analysis of production LLM systems through 10,000+ multi-turn adversarial interactions, discovering critical prompt injection vulnerabilities
- Architected novel context-aware defense framework addressing fundamental limitations in existing industry solutions, researching new security paradigm for enterprise conversational AI systems

PhD Research Assistant at Clemson University 2024 - Present

- Published 4 first-authored papers advancing explainable AI for autonomous systems, with research presented at SAE WCX 2025 and USENIX VehicleSec 2025
- Conducted collaborative research across Army VIPR program and BMW partnership, bridging academic research with real-world deployment in safety-critical and production AI systems
- Mentored undergraduate students on research projects in explainable AI and VLMs

Lead Engineer & Technical Architect at MUSEOLAB 2016 - 2021

- Led end-to-end development of 6 enterprise mobile applications with integrated ML capabilities
- Architected and deployed 5 production ML pipelines from scratch using PyTorch, scikit-learn, NumPy, and Pandas, implementing supervised and unsupervised models (Random Forests, XGBoost, K-Means, PCA, SVD) for business intelligence and real-time inference
- Designed scalable backend infrastructure including RESTful APIs with Flask and SQLAlchemy, web scraping systems processing 150K+ products, and SQL-based inventory management applications
- Drove agile development practices using Scrum to deliver projects and meet client requirements

Publications

Bauerfeind, Philipp et al. (2025). *David vs. Goliath: A comparative study of different-sized LLMs for code generation in the domain of automotive scenario generation*. arXiv: 2510.14115 [cs.SE]. URL: https://arxiv.org/abs/2510.14115.

Duffy, Edward B. et al. (Dec. 2025). "Small Language Models on the Edge for Real-World Agentic Systems in Industry". In: *Southern African Conference for Artificial Intelligence Research (SACAIR)*. To appear. Cape Town, South Africa. URL: https://2025.sacair.org.za/.

Fernandez, David, Pedram MohajerAnsari, Cigdem Kokenoz, Amir Salarpour, Bing Li, and Mert D. Pese (2025). *From MIRAGE to CLEAR: Component-Level Explainable Anomaly Reasoning for Autonomous Vehicle Perception Systems*. Submitted for publication. Manuscript under review.

Fernandez, David, Pedram MohajerAnsari, Cigdem Kokenoz, Amir Salarpour, Bing Li, and Mert D. Pesé (Aug. 2025). "WIP: From Detection to Explanation: Using LLMs for Adversarial Scenario Analysis in Vehicles". In: *3rd USENIX Symposium on Vehicle Security and Privacy (VehicleSec '25)*. Seattle, WA, USA: USENIX Association. URL: https://www.usenix.org/conference/vehiclesec25/presentation/fernandez.

Fernandez, David, Pedram MohajerAnsari, Amir Salarpour, and Mert D. Pese (2025). *Meta-VLM: Evaluating Metadata for VLM Driving Decisions*. Submitted for publication. Manuscript under review.

Fernandez, David, Pedram MohajerAnsari, Amir Salarpour, and Mert D. Pesé (Apr. 2025). "Avoiding the Crash: A Vision-Language Model Evaluation of Critical Traffic Scenarios". In: SAE Technical Paper 2025-01-8213. DOI: 10.4271/2025-01-8213. URL: https://doi.org/10.4271/2025-01-8213.

MohajerAnsari, Pedram et al. (June 2025). "Attention-Aware Temporal Adversarial Shadows on Traffic Sign Sequences". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 3600–3608.

Skills

Programming Languages	Python, Java, Swift, JavaScript, Shell Scripting, HTML
Libraries & Frameworks	NumPy, Pandas, PyTorch, CUDA, Scikit-learn, Matplotlib, D3, Flask