JEFFERSON ENRIQUE HERNANDEZ CEVALLOS

Houston, Texas | Phone: (+1) (832) 758-6250
jefehern@rice.edu, jeffhernandez1995.github.io

SUMMARY

Ph.D. student at Rice University advised by Vicente Ordóñez-Román, researching multimodal LLMs, evaluation for generative models, and self-supervised learning for images/videos. Recent work spans: **Panel-of-Peers** (ICCV'25; iterative peer-graded alignment for LVLMs, +19% across 15 benchmarks), **GenLLaVA** (single-stage instruction tuning for understanding, generation, and editing), **ViC-MAE** (ECCV'24; contrastive masked pre-training across images & videos), **cFreD** (conditional Fréchet Distance for T2I/T2V), and **GViT** (Gaussian image representations with gradient-guided allocation). Experienced in end-to-end systems, datasets, training, evaluation, and deployment, with collaborations at Adobe Research, Google DeepMind, and Meta Reality Labs.

RESEARCH AND WORK EXPERIENCE

Rice University — **Vision, Language and Learning Lab (vislang)**
Ph.D. student — *Aug 2022 to present*

- **GenLLaVA**: Unified LVLM for understanding, T2I generation, and image editing via single-stage generative instruction tuning; preserved comprehension while adding strong generation/editing.
- **ViC-MAE (ECCV 2024)**: Unified self-supervised pretraining across images & short videos combining masked reconstruction with pooled-feature contrastive learning; improved transfer to classification, detection and segmentation.
- **cFreD (WACV 2026)**: Conditional Fréchet Distance that jointly captures realism and text faithfulness for T2I/T2V; stronger preference correlation and robustness than FID+CLIP combinations.
- **GViT**: Mid-level Gaussian image representation with gradient-guided allocation for ViTs; competitive accuracy and interpretability vs. patchified pixels.
- Contributed video support to `ffcv` library for 2–5× faster loading vs. standard PyTorch dataloaders.
- Support: Google CURe & TPU Research programs.

Meta Inc. — **Reality Labs**
Research Intern — *May 2025 to Aug 2025*

- Built/iterated a pipeline that improves **LLM spatial reasoning** to synthesize **binaural audio** from monaural audiobooks by predicting scene/character positions and room acoustics.
- Designed evaluation rubrics (LLM-as-a-judge) and heuristics (rule-based scoring) for spatial coherence and dialogue attribution; produced listener demos and internal tooling.
- Designed a human-in-the-loop data-collection & evaluation pipeline to validate that automatic metrics track human judgments.
- Trained models with GRPO using a custom, domain-specific reward that blends deterministic checks (quote attribution accuracy, alias resolution), differentiable proxies (position smoothness/continuity), and penalty terms (speaker overlaps/crossings); ablated reward weights and sampling strategies to stabilize learning.

Adobe Inc. — **Adobe Research**
Research Intern — *May 2024 to Aug 2024*

- Proposed and implemented **Panel-of-Peers (PoP) (ICCV 2025)**: a label-free, peer-graded self-improvement loop for LVLMs using synthetic preferences and **SimPO/DPO/Best-of-N** style tuning.
- Achieved **19%** average increase across **15** VLM benchmarks (e.g., MMbench, SEED-B, MM-Vet, ScienceQA), and showed knowledge transfer (weak OCR model acquiring OCR skill from peers).
- Delivered ablations on pointwise vs. ranking rewards, uncertainty-weighted aggregation, and SFT vs. SimPO at scale.

AdaViv — **AdaViV**
Computer Vision Engineer — *Dec 2022 to Aug 2022*

- Owned the end-to-end CV pipeline on AWS; trained and deployed plant-part segmentation at scale.
- Train and deploy image segmentation models for plant part segmentation.
- **10x** cost reduction for inference (**$5 → $0.50 per 1K images**) via batching, model/IO optimizations, and autoscaling.

- Built pseudo-labeling loops and data curation to reduce labeling spend; managed labeling vendors and quality assurance.

Escuela Superior Politécnica del Litoral **Industrial Artificial Intelligence (INARI) Research Lab**
Research Assistant *Nov 2018 to Dec 2022*

- Research on unsupervised feature extraction for sequential data using a modified, restricted Boltzmann machine model. This project resulted in two publications.
- Worked on the design of a fast Multi-Object Tracking System for video, which is currently used by a retail brand in Ecuador, one of the laboratory funding companies. This project resulted in two publications.
- Took part in a project on Human Action Recognition through Pose estimation. The research was aimed towards robustness to camera angle variations.
- Led project to estimate conversion rates of retail stores using videos. This project can help estimate how many people who enter a store actually buy a product.
- Led project to calculate customer heatmaps from videos in a retail store. This project is used to better understand customer patterns and improve the layout of store facilities.

Escuela Superior Politécnica del Litoral **Industrial Engineering Bachelor**
Student *Apr 2014 to Oct 2019*

- Led a project to design a program that generates pricing policies and bus fleet schedules using integer linear programming for the university bus company (still used today). I developed a web interface for the program using Python and Django.
- Design a program that optimizes a congested traffic intersection in the university using discrete event-based simulation. I used the simulation software SUMO and some C++.
- My thesis was about measuring manual labor performance automatically using cameras. This project will be used to assess operational decisions and to understand the origin of time losses caused by delays and their impact on the activities in manual tasks. This project resulted in two publications.

PUBLICATIONS

- **Hernandez, J.**, Shi, J., Jenni, S., Ordonez, V., & Kafle, K. (2024). Improving Large Vision and Language Models by Learning from a Panel of Peers. ICCV 2025.
- **Hernandez, J.**, He, R., Balakrishnan, G., Berg, A. C., & Ordonez, V. (2025). GViT: Representing Images as Gaussians for Visual Recognition. arXiv preprint arXiv:2506.23532. Submitted to NeurIPS 2025.
- Xiao, Z., Koo, J., Ouyang, S., **Hernandez, J.**, Meng, Y., & Ordonez, V. (2025). ProxyThinker: Test-Time Guidance through Small Visual Reasoners. arXiv preprint arXiv:2505.24872. Submitted to NeurIPS 2025.
- **Hernandez, J.**, Koo, J., Haji-Ali, M., Yang, Z., & Ordonez, V. (2025). Evaluating Text-to-Image Synthesis with a Conditional Fréchet Distance. arXiv preprint arXiv:2503.21721. Submitted to WACV 2026.
- **Hernandez, J.**, Villegas, R., & Ordonez, V. (2024). Generative Visual Instruction Tuning. arXiv preprint arXiv:2406.11262.
- **Hernandez, J.**, Villegas, R., & Ordonez, V. (2023). ViC-MAE: Self-Supervised Representation Learning from Images and Video with Contrastive Masked Autoencoders. ECCV 2024.
- Palacios, R., Piguave, B. V., **Hernandez, J.**, & Abad, A. G. (2023, October). Automatic Retail Dataset Creation with Multiple Sources of Information Synchronization. In 2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.
- Kim, J. W., **Hernandez, J.**, Cobos, R., Palacios, R., & Abad, A. G. (2022, May). A View Invariant Human Action Recognition System for Noisy Inputs. In 2022 19th Conference on Robots and Vision (CRV) (pp. 67-74). IEEE.
- **Hernandez, J.**, Valarezo, G., Cobos, R., Kim, J. W., Palacios, R., & Abad, A. G. (2021). Hierarchical Human Action Recognition to Measure the Performance of Manual Labor. IEEE Access, 9, 103110-103119.
- **Hernandez, J.**, Lopez, S., Valarezo, G., & Abad, A. G. (2021). Automatic Time and Motion Study Using Deep Learning. In Cyber-Physical, IoT, and Autonomous Systems in Industry 4.0 (pp. 147-162). CRC Press.
- Cobos, R.,**Hernandez, J.**, & Abad, A. G. (2019, June). A fast multi-object tracking system using an object detector ensemble. In 2019 IEEE Colombian Conference on Applications in Computational Intelligence (ColCACI) (pp. 1-5). IEEE.
- Cobos R., **Hernandez, J.**, Abad A.G. (2019) Retail Traffic-Flow Analysis Using a Fast Multi-object Detection and Tracking System. In: Orjuela-Cañón A., Figueroa-García J., Arias-Londoño J. (eds) Applications of Computational Intelligence. ColCACI 2019. Communications in Computer and Information Science, vol 1096.

Springer, Cham.

- **Hernandez, J.**, & Abad, A. G. (2018, May). Learning from multivariate discrete sequential data using a restricted Boltzmann machine model. In 2018 IEEE 1st Colombian Conference on Applications in Computational Intelligence (ColCACI) (pp. 1-6). IEEE.
- **Hernandez, J.**, & Abad, A. G. (2018, May). Spatial and Temporal Feature Extraction Using a Restricted Boltzmann Machine Model. In IEEE Colombian Conference on Applications in Computational Intelligence (pp. 3-13). Springer, Cham.

EDUCATION

- Ph.D student, Computer Science, Rice University, 2022-present.
- B.Sc., Industrial Engineering, Escuela Superior Politécnica del Litoral, 2014-2019.

TECHNICAL SKILLS

- **Modeling/Training**: PyTorch, JAX, FSDP, vLLM, Megatron-LM, DeepSpeed, Verl.
- **Vision/Multimodal**: CLIP/SigLIP, Self-Supervised models, LLaVA, Diffusion Models, LLMs.
- **Infra/Perf**: SLURM, Docker, Run:AI, CUDA, Triton kernels, AWS/GCP, model profiling.
- **Modeling:** Statistical and mathematical systems, Markov decision processes, industrial processes.
- **Optimization:** Linear programming, Integer programming, genetic algorithms.
- **Data Analysis:** Probabilistic and Statistical Analysis, Big Data, Database Management.
- **Project Management:** Budget planning, Project planning, and Task management.
- **Programming:** Python, R, MATLAB, GAMS.

TEACHING EXPERIENCE

- 2023 - Teaching Assistant for Deep Learning for Vision and Language.
- 2022 - Teaching Assistant for Deep Learning for Vision and Language.
- 2018 - Teaching Assistant for Integer Linear Programming.
- 2017 - Teaching Assistant for Operations Research.
- 2017 - Teaching Assistant for Advanced Statistics.
- 2016 - Teaching Assistant for Differential Equations.
- 2016 - Teaching Assistant for Multivariate Calculus.

COMMUNITY SERVICE

- 2025 - Served as a mentor for SRI at Harvard's Undergraduate OpenBio Laboratory.
- 2023 - Served as a mentor for the TaReCDa 2023 conference hosted in Machala, Ecuador.
- 2023 - Served as a mentor for the CVPR 2023 LXAI mentorship program.
- 2023 - Served as president of the Latin American Graduate Student Association at Rice University.
- 2019 - Served as a group mentor in a Machine Learning Hackathon. The group I advised ended third in the competition.
- 2019 - Served as a group mentor and challenge designer for my university programming Hackathon. The group I advised won the competition.
- 2018 - Gave free mentorship to students at my university. The topics included programming, calculus, and linear algebra.
- 2017 - Security process re-design for a children's hospital.
- 2014 - Volunteer at a nursing home and at a kindergarten.

OTHER SKILLS

Software: LaTeX, MySQL, Django, Word, Excel, and PowerPoint, Minitab, AutoCAD, Inventor.
Languages: English: professional proficiency (104 TOEFL). Spanish: native.