Ricardo Ignacio Pizarro Carreño

Efficient Video Transformers · Sparse Computation · GPU Kernel Development

Madrid, Spain | +34 611288286 | ricardo1459@gmail.com | github.com/RicardoP0

Featured Research

SV-TAD: Native Sparse Convolutions for Efficient Multitask Video Understanding (ECCV 2026)

- Designed **native sparse 2D convolutions**, a new computational primitive that lets convolutional ViT adapters operate directly on dynamically pruned token sequences via precomputed *neighbor-index tables*, removing the dense-grid reconstruction step that previously bottlenecked token-sparse adapters.
- Implemented custom forward/backward **CUDA kernels** (cuBLAS implicit-GEMM for large channel widths plus fused tiled kernels for smaller ones) and released them as a **standalone sparse-convolution library for PyTorch**, reaching up to 8× kernel speedup at low keep rates.
- Combined this primitive with attention-based, **FlashAttention**-compatible token selection inside a sparse bottleneck adapter (frozen VideoMAEv2 / InternVideo backbones), cutting backbone compute by up to **64%** and giving **2.2**× faster inference and **1.7**× faster training, while matching SOTA mAP on THUMOS-14 and ActivityNet-1.3.
- Extended the adapter with **cross-attention auxiliary tokens** for multitask learning, enabling auxiliary landmark/pose supervision that improves fine-grained assembly action detection on ATTACH by **+2.41% mAP**.

Education

PhD in Electronics: Advanced Electronic Systems. Intelligent Systems **May 2023 – Oct. 2026 (Expected)**
Universidad de Alcalá *Madrid, Spain*
– Thesis: Multi-view and landmark augmented deep learning model for HMI in autonomous vehicles

MSc. Computer Science **March 2020 – August 2021**
Universidad Católica del Norte *Antofagasta, Chile*
– Thesis: Multimodal fusion technique for video sequences
– Graduated with highest honors (*Máxima distinción*)

Engineer's Degree in Computer Science and Informatics **March 2014 – August 2021**
Universidad Católica del Norte *Antofagasta, Chile*
– Graduated with honors (*Con distinción*)

Technical Skills

Programming & GPU Systems: Python, C++, CUDA — custom CUDA kernels and PyTorch C++/CUDA extensions for sparse and efficient computation
Deep Learning: PyTorch, PyTorch Lightning, FlashAttention, scikit-learn, Weights & Biases
Research Areas: Vision Transformers; efficient & sparse computation (token pruning/merging, native sparse convolutions, adapter-based PEFT); temporal action detection/localization; multitask learning; human action recognition
Languages: English – Full professional proficiency, Spanish – Native, Japanese – Intermediate, German – Beginner

Research and Teaching Experience

PhD Researcher **May 2023 – Present**
Universidad de Alcalá

Advisor: Dr. Luis M. Bergasa, Dr. Luis Baumela
– Researching efficient video transformer architectures for human action recognition and temporal action detection in low-compute environments, combining token selection, native sparse computation, and custom GPU kernels (see *Featured Research* above).
– Established further SOTA results across multiple domains, including:
  * Driver Distraction (Drive&Act, 3MDAD, 100-Driver)
  * Activities of Daily Living (Toyota Smarthome, NTU RGB+D 120)
  * Temporal Action Localization (THUMOS14, ATTACH)

Multimodal Sharp Wave Ripple Detection **Aug. 2022 – Feb. 2023**
Donders Institute for Brain, Cognition and Behavior

Research Intern | Supervisor: Dr. Lisa Genzel
- Ensured dataset integrity by implementing robust data labeling workflows and diagnostic tools to identify and correct annotation errors.
- Collaborated with a team of neuroscientists to perform feature engineering and analyze feature importance, directly improving classifier performance.
- Built a flexible deep learning framework using PyTorch Lightning and Weights & Biases (Wandb) to systematically benchmark CNN and Transformer architectures, enabling robust experiment tracking and hyperparameter optimization.

Multimodal Fusion Technique for Video Sequences **March 2020 – June 2021**

Universidad Católica del Norte

MSc. Thesis | Advisor: Dr. Juan Bekios-Calfa
- Modified a multimodal fusion technique to leverage spatio-temporal relations in a video while using self-supervised learning.
- Leveraged visualization techniques to validate the hypothesis of the proposed fusion technique.
- Achieved an improvement of 2.7 points over the baseline and obtained state-of-the-art results.

Image Segmentation in Mineral Flotation **July 2020 – October 2020**

Universidad Católica del Norte

Student Researcher | Computer Science Department | PI: Dr. Juan Bekios
- Designed experiments to evaluate different vision models for a novel application in mineral flotation devices.
- Developed a deployment-ready model for in-site testing and a performance tracking pipeline to verify its performance in the real world.

Reinforcement Learning Applied to Wave Energy **April 2019 – July 2019**

Universidad Católica del Norte

Student Researcher | Civil Engineering Department | PI: Roberto Cortés
- Conducted experiments using reinforcement learning algorithms to control a wave energy converter in a wave simulator.
- Worked alongside a team of civil engineers to assess the performance of different algorithms, comparing them to theoretical results while finding possible sources of inefficiency.

Predictive Model for On-Time Graduation Projects **July 2018 – May 2019**

Universidad Católica del Norte

Student Researcher | Computer Science Department | PI: Brian Keith
- Created machine learning models to predict whether a student will be late in submitting their thesis, and statistical analysis to test the significance of these results.
- Leveraged feature importance techniques to find factors that influenced the prediction. The resulting model was shared with the department, and a plan was created to test the model in the real world.

Teaching Assistant **February 2016 – June 2020**

Universidad Católica del Norte

Courses: Data Structures, Software Engineering, Artificial Intelligence, and Intelligent Systems (M.Sc. level)
- Designed and led weekly lectures for groups of ~40 students.
- Updated and translated course materials, including study problems, projects, and exams.
- Graded problem sets and exams.

Teaching Assistant **June 2020 – December 2020**

Universidad Adolfo Ibáñez
- Planned deep learning projects related to image/video classification, NLP, and image segmentation.
- Created and translated course materials.

Professional Experience

Founder & Chief Technology Officer **February 2019 – Present**

PIGNUS
- Co-founded Pignus and led development of **Rehaviour**, a VR-based behavioral safety platform that uses gamified simulations to assess and improve worker safety behavior in high-risk industries (mining, construction, energy, logistics), deployed with enterprise clients including Glencore, Copec, Equans, Albemarle, and Finning.
- Designed, deployed, and maintained the company's **multi-cloud infrastructure** across AWS, Azure, and GCP, covering the VR application backend, a customer-facing web reporting platform built with **PHP/Laravel** and **MySQL**, data storage, and CI/CD pipelines.

– Designed and trained the platform's **AI/ML models**, including unsupervised models that analyze in-VR behavioral data to identify safety-competency gaps; collaborated with a team of psychologists to relate user actions to behavioral and cognitive profiles.
– Built end-to-end **data pipelines** on **Firestore** and event-driven **Cloud Run** jobs (Python) to process raw VR session data at scale, feeding model training and a 24/7 analytics dashboard that turns it into objective safety reports for client organizations.
– Led cross-functional development teams (VR, web, ML, infrastructure) and owned technical feasibility assessments and roadmap planning for AI and VR initiatives.

Deep Learning Engineer **February 2021 – June 2021**

SERCOL

– Created experiments to benchmark several deep learning architectures for OCR on tables, and developed algorithms to format the semi-structured results obtained from OCR models.
– Designed and developed a scalable data pipeline to handle different types of files, achieving the automation of document processing and cutting the processing time by almost 75%.

Relevant Courses

Neuromatch Academy **July 2022**

Biologically inspired neural networks and stochastic processes

Computational Neuroscience **October 2021**

University of Washington

Artificial Intelligence and Big Data Techniques **January 2019 – February 2019**

Universidad Politécnica de Madrid

Publications

- **Pizarro, R.**, Valle, R., Buenaposada, J. M., Bergasa, L. M., & Baumela, L. (2026). SV-TAD: Native Sparse Convolutions for Efficient Multitask Video Understanding. *ECCV*.
- **Pizarro, R.**, Valle, R., Buenaposada, J. M., Bergasa, L. M., & Baumela, L. (2026). ADA-SEL: Efficient Training and Inference for Video TAD. *Image and Vision Computing*. Under review.
- **Pizarro, R.**, Barea, R., Buenaposada, J. M., Baumela, L., & Bergasa, L. M. (2026). Temporal Action Localization for Driver Action Recognition: A Benchmark with Efficient Token Selection. *IEEE Intelligent Transportation Systems Conference (ITSC 2026)*.
- **Pizarro, R.**, Valle, R., Barea, R., Buenaposada, J. M., Baumela, L., & Bergasa, L. M. (2026). PO-GUISE+: Pose and Object Guided Transformer Token Selection for Efficient Driver Action Recognition. *IEEE Transactions on Intelligent Transportation Systems*. doi: https://doi.org/10.1109/TITS.2026.3665254.
- **Pizarro, R.**, Valle, R., Buenaposada, J. M., Bergasa, L. M., & Baumela, L. (2025). Pose-guided token selection for the recognition of activities of daily living. *Image and Vision Computing*, 162, 105686. https://doi.org/10.1016/j.imavis.2025.105686.
- **Pizarro, R.**, Bergasa, L. M., Baumela, L., Buenaposada, J. M., & Barea, R. (2024). DRVMon-VM: Distracted driver recognition using large pre-trained video transformers. In *Proceedings of the 2024 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1901–1906). IEEE. https://doi.org/10.1109/IV55156.2024.10588412.
- Quelopana, A., Keith, B., & **Pizarro, R.** (2024). Predictive modeling of on-time graduation in computing engineering programs: A case study from Northern Chile. *Computer Applications in Engineering Education*, e22767. https://doi.org/10.1002/cae.22767.
- **Pizarro, R.**, & Bekios-Calfa, J. (2021). Emotion recognition using multimodal matchmap fusion and multi-task learning. *11th International Conference on Pattern Recognition Systems (ICPRS 2021)*.

Presentations

- Emotion recognition using multimodal matchmap fusion and multi-task learning. Online presentation, *11th International Conference on Pattern Recognition Systems (ICPRS 2021)*.
- Estimación del estado emocional de una persona a través de secuencias de video multimodales utilizando técnicas de aprendizaje automático multitask y Deep Learning. *INFONOR/WDKE'2020*.
- Sentiment Analysis of Paper Reviews based on Deep Learning Neural Networks. *INFONOR/WDKE'2020*.