Diego Revilla

diego@diegorevilla.ai linkedin.com/in/diego-ai — github.com/diegorevillar — diegorevilla.ai

Experience

HiTZ Center for Language Technology, U. Basque Country San Sebastian, Spain
Student Researcher — Multimodal Learning **Oct 2026 — May 2027**

Apple Seattle, United States
Machine Learning Engineer Intern — Diffusion Models **Jun 2026 — Sep 2026**

Developed a diffusion model that estimates the complete floorplan of a house by observing a partial set of disjoint rooms, achieving 89% floorplan reconstruction accuracy when 50% of rooms are observed.

- Constructed a dual-token training pipeline for the indoor place recognition task that surpasses modern architectures such as AnyLoc & frozen DINOv3 by 12% Recall@3 on our internal dataset.
- Trained a DQN active user querying policy that exploits the model's uncertainty to choose what to ask, gaining an extra 3% accuracy over a greedy policy, while keeping prompts to less than 3.
- Built a 72 GB dataset of 10k home floorplans expressed as heterogeneous graphs, including room and pairwise connectivity semantics, classification, meshes and room-level panoramic cubemaps.

HiTZ Center for Language Technology, U. Basque Country San Sebastian, Spain
Student Researcher Fellow — Multimodal Learning **Sep 2025 — May 2026**

Designed and implemented a stateful reasoning architecture to iteratively solve spatially compositional VQA tasks by applying latent recursive reasoning in vision-language models without text-based CoT.

- Introduced conditioned attention shifting, which makes the model attend to different input regions at each reasoning step, reaching 80% F1 on VSR & GQA-rel (+6% over single-pass).
- Trained a specialist model that surpasses off-the-shelf foundation models InternVL2-8B & Qwen2.5-VL-7B on the same datasets with 400M parameters on relative positional and spatial tasks.
- Parallelized VLM training over 4 H100 GPUs using DDP, shrinking training time by 70%.

National University of Singapore Singapore
Research Intern — Computer Vision & Neural Rendering **Jun 2025 — Sep 2025**

Conceived a progressive neural compression codec for Gaussian Splatting to enable neural rendering on bandwidth-constrained client-server infrastructures, allowing for first render times under 3 seconds.

- Reduced the total file size by modeling 3D spatial context to exploit local geometry and feature similarities to reduce uncertainty entropy, obtaining a 21% compression gain compared to HAC.
- Matched current methods in perceptual, structural and quality metrics (>27dB PSNR, >0.87 SSIM, <0.21 LPIPS) despite decreased size on the BungeeNeRF & MipNeRF 360 datasets.

Deusto Institute of Technology - *Research Assistant (NLP)* **Sep 2024 — May 2025**
Sociograph Neuromarketing - *Data Science Intern* **Jun 2024 — Aug 2024**
Deusto Institute of Technology - *Research Assistant (VLMs)* **Oct 2023 — May 2024**
On-Line Technical Solutions - *Software Developer Intern* **Jun 2023 — Aug 2023**

Education

University of Deusto, Bilbao, Spain **Sep 2022 — Jun 2027**
BSc. in Data Science & Artificial Intelligence *Distinction in 10 courses*
BEng. in Computer Engineering *Distinction in 8 courses*
5-year Double Degree Program, Student Representative, Merit Scholarship GPA: 3.8/4 (Top 2%)

Skills

Programming Languages: Python, C, C++, Java, Swift, Scala, R, HTML, SQL, LaTeX, x86 ASM
Stack: PyTorch, TensorFlow, CoreML, Scikit-learn, NumPy, Pandas, W&B, Docker, Nsight, OpenGL, Git

Selected Publications

Revilla, D., Fernandez-de-Retana, M., Chen, L., Bilbao-Jayo, A., & Fernandez-de-Retana, M. "A Counterfactual Graph Approach to Oppositional Narrative Analysis." *Trustworthy AI for Good Workshop (AI4GOOD @ NeurIPS 2026).*

Revilla, D., Suresh, P., Bhojan, A., & Tsang, O. W. "SCAR-GS: Spatial Context Attention for Residuals in Progressive Gaussian Splatting." *The 5th Workshop on Computer Vision for Multimedia Spatial Intelligence through Time, 2026 (MUSTCV @ ECCV 2026).*