

Ayan Banerjee

[ayanban011](#) | [ayan-banerjee-733b5b16b](#) | [ayanban011.github.io](#) | abanerjee@cvc.uab.es | [+34-677640371](#) | [AyanBanerjee](#)

Education

PhD in Computer Science

March, 2024 - Present

Universitat Autònoma de Barcelona, Spain

- Research on Layout-Guided Diffusion Models and Mechanistic Interpretability.

M.S. in Computer Vision

Oct, 2022 - Sep, 2023

Universitat Autònoma de Barcelona, Spain

- Research on Document Layout Analysis in Self-Supervised Settings. *CGPA:8.08*

B.Tech in Computer Science & Engineering

Aug, 2017 - July, 2021

Jalpaiguri Government Engineering College, India

- Research on Scene Text Recognition. *CGPA:9.15*

EXPERIENCE

Sony AI, Barcelona

Sep, 2026 - Present

Research Intern (Supervised by Joan Serra)

- Research on self-supervised music matching.

People-Centred AI, University of Surrey, UK

Jan, 2025 - Feb, 2025

Research Intern (Supervised by Prof. Anjan Dutta)

- Research on End-to-End Comics Generation with Diffusion Models.

Computer Vision Center, Spain

Jun, 2022 - Sep, 2022

Research Intern (Supervised by Prof. Josep Lladós)

- Research on Graph-based Knowledge Distillation towards Document Layout Analysis.

Cognizant Technology Solutions, India

Aug, 2021 - May, 2022

Program Analyst

- Develop an API for an Automated Pension Management System.

PUBLICATIONS

- **Banerjee A**, Chaudhuri A, Lladós J, Pal U, Dutta A. Can Circuit Alignment Predict OOD Generalization?. (**NeurIPS 2026**)
- Mondal A, Nag S, **Banerjee A**, Lladós J, Zhu X, Dutta A. CountLoop: Training-Free High-Instance Image Generation via Iterative Agent Guidance. (**TMLR 2026**)
- **Banerjee, A.**, Mathur, N., Lladós, J., Pal, U. and Dutta, A., 2026, March. CraftSVG: Multi-Object Text-to-SVG Synthesis via Layout Guided Diffusion. (**WACV 2026**)
- Purkayastha, K., **Banerjee, A.**, Lladós, J. and Pal, U., 2026. DocRevive: A Unified Pipeline for Document Text Restoration. (**CVPR-W 2026**)
- **Banerjee, A.**, Vilariño, F. and Lladós, J., CraftGraffiti: Exploring Human Identity with Custom Graffiti Art via Facial-Preserving Diffusion Models. (**NeurIPS Creative AI track, 2025**)

- **Banerjee A**, Biswas S, Lladós J, Pal U. GraphKD: Exploring knowledge distillation towards document object detection with structured graph creation. (**ICDAR, 2024**)
- **Banerjee A**, Biswas S, Lladós J, Pal U. SemiDocSeg: harnessing semi-supervised learning for document layout analysis. (**IJDAR 2024**)
- Maity S, Biswas S, Manna S, **Banerjee A**, Lladós J, Bhattacharya S, Pal U. Selfdocseg: A self-supervised vision-based approach towards document segmentation. (**ICDAR, 2024**)
- **Banerjee A**, Biswas S, Lladós J, Pal U. Swindocsegmenter: An end-to-end unified domain adaptive transformer for document instance segmentation. (**ICDAR 2024**)
- Das A, Biswas S, **Banerjee A**, Lladós J, Pal U, Bhattacharya S. Harnessing the power of multi-lingual datasets for pre-training: Towards enhancing text spotting performance. (**WACV 2024**)

PROJECTS

LoRA Merging for Concept Unlearning in Diffusion

April, 2026 - Present

Performing Modular Concept Erasing by treating unlearning as a weight-subtraction task. Specifically, we isolate the concept delta from a toy model and subtract it from a foundation model via merging, achieving unlearning without ever fine-tuning the base model. (*Supervised by Dr. Joost van de Weijer, Computer Vision Center, Spain*)

Evolution of Neural Circuit Under Distribution Shift

Aug, 2025 - Present

Developed a graph-kernel-based Circuit Similarity Metric to prove that high model expressivity overcomes initialization randomness by forcing the optimization landscape to form robust, class-specific circuits that resist out-of-distribution shifts. (*Supervised by Dr. Abhra Chaudhuri, People-Centered AI, UK*)

SKILLS

Deep Learning: Pytorch, Tensorflow, Keras; **Diffusion Models:** SD1.5, SDXL, SD3.5, FLUX.1 dev, FLUX.1 Kontext, IP-Adapter; **MultiModal Models:** CLIP, Qwen3.5; **Mechanistic Interpretability:** Attribution Patching, TransformerLens.

ACHIEVEMENTS

- Best Student Paper Award Winner in ICDAR 2024, for GraphKD.
- FI-Joan Oró Ph.D. fellowship, 2024.
- 1st in the Researchathon 2021 healthcare track, organised by the United Nations Environment Programme and the Harvard Kennedy School.
- 1st Runner Up in Nexas-A-Thon in Waste Management Track organised by Indian Institute of Technology Guwahati.

ACADEMIC CONTRIBUTIONS

- **Conference Reviewer:** CVPR, NeurIPS, ICLR, ICML, BMVC, WACV.
- **Journal Reviewer:** Pattern Recognition, IEEE Transaction of Multimedia
- **Conference Attended:** ICDAR 2024, EurIPS 2025, DLBCN 2023 - 25
- **Other Events:** Summer School on Foundation Models (Amsterdam 2024)