

Oriol Barbany Mayor

☎ +34 608 54 45 47 | ✉ obarbany@iri.upc.edu | 🌐 barbany.github.io | 📱 Barbany | 🌐 [oriol-barbany](https://oriol-barbany.com)

Summary

PhD student working on computer vision for deformable object understanding and manipulation. I have followed graduate courses in mathematics and theoretical computer science, with an emphasis on machine learning, and I have a strong background in signal processing. Accustomed to a high level of workload and always eager to learn.

Education

Institut de Robòtica i Informàtica Industrial (IRI), CSIC-UPC

Barcelona

PHD IN ARTIFICIAL INTELLIGENCE - ELLIS PHD & POSTDOC PROGRAM

Fall 2021 - Present

- **Project:** Feedback models for real-time robotic cloth manipulation.
- **Exchange institution (as part of the ELLIS program):** École Polytechnique Fédérale de Lausanne (EPFL).
- **Supervisors:** Prof. Carme Torras and Dr. Adrià Colomé (IRI), and Prof. Amir Zamir (EPFL).

École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne

MSC. IN DATA SCIENCE

2018 - 2020

- **Average score:** 5.74/6.00
- **Activities:** Member of the *Orchestre Symphonique et Universitaire de Lausanne* and the Lausanne Student Chamber Orchestra.
- **Selected coursework:** Advanced Algorithms, Artificial Neural Networks, Automatic Speech Processing, Image Analysis and Pattern Recognition, Information Theory and Signal Processing, Machine Learning, Mathematical Foundations of Signal Processing, Mathematics of Data, Optimization for Machine Learning, Statistics for Data Science, and Time Series.

Universitat Politècnica Catalunya (UPC)

Barcelona

BSC. IN TELECOMMUNICATIONS TECHNOLOGIES AND SERVICES ENGINEERING

2014 - 2018

- **Average score:** 8.49/10.00. 4th of the promotion. See [recognition of outstanding academic achievement](#).
- **Major:** Audiovisual Systems.
- **Course honors:** [Bachelor's thesis](#) [8], Image and Video Processing, Acoustics and Electro-acoustics, Introduction to Deep Learning, Radiation and Propagation, Introduction to Audiovisual Processing, Electromagnetic Waves, Signals and Systems, Electromagnetism, Introduction to Mathematics, and Fundamentals of Computers.

Escola Municipal de Música i Conservatori Professional de Sabadell

Sabadell

PROFESSIONAL DEGREE IN MUSIC

2002-2014

- Bassoon and harmony studies.

Experience

Visual Intelligence and Learning Laboratory @ EPFL

Lausanne

VISITING DOCTORAL STUDENT

May 2024 - Present

- **Project:** "3D Multimodal Models"
- **Summary:** Research on massively masked multimodal modeling for any-to-any prediction of 3D-grounded modalities (e.g., point clouds, voxels, text, images with camera poses, radiance fields, etc.).
- **Supervisor:** Prof. Amir Zamir

Amazon

Palo Alto

APPLIED SCIENTIST INTERN

Summer 2023

- **Project:** "Multimodal Search"
- **Summary:** Research at the Visual Search & Augmented Reality team to include multimodal search functionalities in the current search systems. Improved an internal model, created an internal benchmark for Large Language Models (LLMs) targeting multimodal retrieval, and proposed an approach to adapt Vision-Language Models (VLMs) to the composed retrieval problem [3].
- **Supervisors:** Dr. Arnab Dhua, Dr. Xinliang Zhu, and Michael Huang

Amazon

Palo Alto

APPLIED SCIENTIST INTERN

Summer 2022

- **Project:** "Robust visual similarity learning"
- **Summary:** Worked with the Visual Search & Augmented Reality team to improve the Amazon Lens product (take a picture of anything and find the most similar product on the Amazon catalog). Proposed a solution to learn better visual similarity models (extracting a representation that correctly encodes the semantics of the image) using datasets with many wrong labels [4].
- **Supervisors:** Dr. Arnab Dhua, Dr. Xiaofan Lin, and Dr. Muhammet Bastan

Amazon

Berlin

APPLIED SCIENTIST INTERN

Fall 2020

- **Project:** "Contextual Position Bias Estimation for Unbiased Learning to Rank"
- **Summary:** Devised a ranker-agnostic method to estimate the position bias (the tendency of users to engage with top-ranked results) under different contexts (e.g., user-specific browsing patterns, type of query, or type of device) from observational data (without needing feedback from users after performing interventions). The obtained method obtains better position bias estimates than ranker-dependent and intervention-based algorithms, which leads to an improved ranking performance [5].
- **Supervisor:** Dr. Matteo Ruffini and Dr. Giuseppe di Benedetto.

Logitech Europe S.A.

Lausanne

ENGINEERING INTERN, AI TEAM

Spring 2020

- **Project:** "Speech and Audio Style Transfer" - Internship credited with [Master Thesis](#)
- **Summary:** This project introduced FastVC, a method for fast (4x faster than real-time on CPU) many-to-many (dealing with several speakers as sources and targets with a single model) nonparallel (trained with a dataset with source and target speakers uttering different phrases) Voice Conversion (VC). The proposed approach was accepted as a technical paper for the VC Challenge 2020, where FastVC outperformed the Challenge baselines in quality results on the cross-lingual VC task [6].
- **Supervisors:** Dr. Milos Cernak (Logitech) and Prof. Martin Jaggi (EPFL).

Machine Learning and Optimization Laboratory @ EPFL

Lausanne

STUDENT RESEARCH ASSISTANT

Fall 2019

- **Project:** "Affine-Invariant Robust Training" - Semester project
- **Summary:** Proposed evolution strategies as zeroth order optimization algorithms to build robust models against affine transformation attacks (which, unlike ℓ_p bounded adversarial attacks, covered in most works, are likely to appear in the real world) [7].
- **Supervisors:** Dr. Sebastian Stick, Maksym Andriushchenko (PhD candidate), and Prof. Martin Jaggi.

Publications

- [1] David Blanco-Mulero, **Oriol Barbany**, Gokhan Alcan, Adrià Colomé, Carme Torras, and Ville Kyrki. "Benchmarking the Sim-to-Real Gap in Cloth Manipulation". In: *IEEE Robotics and Automation Letters* 9.3 (2024), pp. 2981–2988.
- [2] **Oriol Barbany**, Adrià Colomé, and Carme Torras. "Deformable Surface Reconstruction via Riemannian Metric Preservation". In: *Computer Vision and Image Understanding* 249 (2024), p. 104155.
- [3] **Oriol Barbany**, Michael Huang, Xinliang Zhu, and Arnab Dhua. "Leveraging Large Language Models for Multi-modal Search". In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2024.
- [4] **Oriol Barbany**, Xiaofan Lin, Muhammet Bastan, and Arnab Dhua. "ProcSim: Proxy-Based Confidence for Robust Similarity Learning". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 2024, pp. 1308–1317.
- [5] **Oriol Barbany Mayor**, Vito Bellini, Alexander Buchholz, Giuseppe Di Benedetto, Diego Marco Granziol, Matteo Ruffini, and Yannik Stein. *Ranker-agnostic Contextual Position Bias Estimation*. arXiv:2107.13327. 2021.
- [6] **Oriol Barbany** and Milos Cernak. "FastVC: Fast Voice Conversion with non-parallel data". In: *Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge*. 2020.
- [7] **Oriol Barbany Mayor**. *Affine-Invariant Robust Training*. arXiv:2010.04216. 2020.
- [8] **Oriol Barbany**, Antonio Bonafonte, and Santiago Pascual. "Multi-Speaker Neural Vocoder". In: *IberSPEECH*. 2018.

Honours and Volunteering

HONORS

- | | | |
|------|---|-----------|
| 2018 | 1st prize , Hackathon Mobility: SEAT Challenge - AI For Mobility & Driving Experience. See prize and certificate | Barcelona |
| 2014 | Honour Prize , Bassoon category: Granted during 6th course. See certificate | Sabadell |
| 2011 | Honour Diploma , Bassoon category: Granted during 3rd course | Sabadell |

VOLUNTEERING

- | | | |
|------|--|------------------|
| 2020 | Volunteer , International Conference on Machine Learning (ICML). See certificate . | Vienna (Virtual) |
| 2018 | Student Volunteer , IEEE Wireless Communications and Networking Conference. See certificate . | Barcelona |

Skills

Deep Learning Frameworks	Tools	Wandb, TensorBoard, Jupyter Notebook, Git/GitHub, AWS, Slurm, RunAI, \LaTeX , Linux
	Python Packages	NumPy, SciPy, scikit-learn, Matplotlib, Seaborn, Hydra, accelerate, peft, transformers
		PyTorch, TensorFlow, JAX
	Languages	Catalan (Native), Spanish (Native), English (IELTS: 6.5 - see certificate , TOEIC: 855 - see certificate), French (A1/A2 - see certificate)