

Ariel Gjaci

Postdoctoral Researcher | Multimodal Generative Learning | Audio-Visual Learning | Human Motion

ariel.gjaci@iit.it | +393483776041 | Genoa, Italy

[Website](#) | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#)

Postdoctoral researcher with a PhD in Artificial Intelligence and Robotics. Develops multimodal generative models for human motion and spatial audio, combining diffusion transformers, representation learning, dataset development, and experimental evaluation in Python and PyTorch.

TECHNICAL SKILLS

Generative Models

Diffusion transformers; GANs;
Multimodal conditioning

Learning and Evaluation

Representation learning; Domain
generalization; Generative model
evaluation

Programming

Python; C/C++; MATLAB/Simulink

Frameworks

PyTorch; Scikit-learn; TensorFlow

Tools and Training

Git; Linux; SLURM; Data parallelism

EDUCATION

Ph.D. Artificial Intelligence and Robotics

University of Genoa

Nov 2021 - May 2025

Research collaboration with King's
College London.

M.Sc. Robotics Engineering

University of Genoa

2018 - 2021

B.Sc. Biomedical Engineering

University of Genoa

2014 - 2018

LANGUAGES

Italian (native)

Albanian (native)

English (fluent)

PROFESSIONAL EXPERIENCE

Postdoctoral Researcher

Italian Institute of Technology (IIT)

Feb 2025 - Present

- Developing few-shot room impulse response generation from audio and depth, with counterfactual training for reverberation control (ongoing research).
- Built Eigenmike 64 processing pipelines for beamforming, binaural rendering, and acoustic mapping on 3D room point clouds.
- Implemented audio/video synchronization and 360-degree acoustic overlays for spatial-audio recordings.

PhD Researcher

University of Genoa

Nov 2021 - May 2025

- Designed ALaDiT, a multimodal diffusion transformer generating 4 s of 3D motion in <14 ms on a single RTX 3090 (SICAGE, ECCV 2026).
- Created and released TED4C-L: 106.45 hours from 764 speakers, aligning MMPose-extracted motion with audio and text.
- Analyzed cultural cues in interaction and affective data using classical and deep learning models (AIRO 2022; AHRI at ACII 2024).
- Benchmarked generative models with controlled ablations; released training and evaluation code.

Software Engineer

AKKA Technologies

May 2021 - Oct 2021

- Developed Python backend software for aircraft panels and presented a working demo to stakeholders.

SELECTED PUBLICATIONS

SICAGE: Speaker-Independent Culture-Aware Gesture Generation using TED4C-L Dataset. European Conference on Computer Vision (ECCV), 2026.

Labeling Sentences with Symbolic and Deictic Gestures via Semantic Similarity. IEEE RO-MAN, 2024.

Towards Culture-Aware Co-Speech Gestures for Social Robots. International Journal of Social Robotics, 2022.

[Full publication list](#)

TEACHING AND SERVICE

Teaching assistant (C++, Android/Kotlin); ECCV workshop co-organizer; reviewer for AAAI, ECCV, ACM Multimedia, IROS, ICMI, and IJSR; IEEE RO-MAN 2024 session chair.