

Lorenzo Bianchi

PHD STUDENT IN VISION-LANGUAGE ARTIFICIAL INTELLIGENCE

Via Panoramica 17B, Pescia (PT), 51012, Italy

+39 331 9744874 | bianchi.lorenzo998@gmail.com | lorebianchi98.github.io | lorenbianchi98 | lorenzo-bianchi-893bb225a | Lorenzo Bianchi

I am a passionate Ph.D. candidate in Vision-Language Artificial Intelligence at the University of Pisa and an Associate Researcher at ISTI-CNR.

My work has been published at top-tier venues such as CVPR (Highlight Poster), ICCV, and WACV, and I have received Best Paper Awards at CBMI and IEEE-CH. In 2025, I joined Disney Research | Studios in Zürich as a Research Intern, where I worked on large-scale diffusion models for text-conditioned image generation. My research interests focus on Vision-Language Models, particularly in understanding, improving, and exploiting the representations learned by foundational models such as CLIP and DINO for image-text tasks.

Education

Ph.D. in Vision-Language Artificial Intelligence

Pisa, Italy

UNIVERSITY OF PISA

Mar. 2023 – Present

- Research interests: Vision-Language AI, Representation Learning, and Unsupervised Learning.
- Supervisors: Fabrizio Falchi, Dr. Fabio Carrara, Dr. Nicola Messina, Dr. Giuseppe Amato.

M.Sc. in Artificial Intelligence and Data Engineering

Pisa, Italy

UNIVERSITY OF PISA

Sep. 2020 – Feb. 2023

- Final Mark: **110/110**.
- Thesis Title: "Design and development of cross-modal retrieval techniques based on transformer architectures".

B.Sc. in Computer Engineering

Pisa, Italy

UNIVERSITY OF PISA

Sep. 2017 – Sep. 2020

- Final Mark: **110/110**.
- Thesis Title: "Dynamics analysis of sports events on online social networks through clusters".

Work Experience

Associate Researcher

Pisa, Italy

CNR (NATIONAL RESEARCH COUNCIL OF ITALY)

Mar. 2023 – Present

- Research Associate in deep learning for vision-language tasks at the Artificial Intelligence for Media and Humanities (AIMH) Lab within the Institute of Information Science and Technologies "Alessandro Faedo" (ISTI).
- Focused on studying, enhancing, and exploiting foundational model representations (CLIP, DINO, etc.) and applying them to a variety of tasks.
- Worked on image-text matching, retrieval, open-vocabulary object detection, open-vocabulary semantic segmentation, and image captioning.
- Collaborated in the Sun XR European project on extended reality and in the ITSERR project, advancing AI applications in Religious Studies.

Research Intern

Zürich, Switzerland

DISNEY RESEARCH | STUDIOS

May 2025 – Aug. 2025

- Designed, implemented, and trained large-scale diffusion models for image generation.
- Supervised by Dr. Vinicius C. Azevedo.

Machine Learning Engineer

Pisa, Italy

WEVO S.R.L.

Apr. 2024 – Jul. 2024

- Developed a sentiment analysis system for social media posts, implementing deep learning models for automatic content classification.

Publications

CountingDINO: A Training-free Pipeline for Class-Agnostic Counting using Unsupervised Backbones

WACV 2026

Tucson, Arizona

G. Pacini*, L. Bianchi*, L. Ciampi, N. Messina, G. Amato, F. Falchi
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)

Talking to DINO: Bridging Self-Supervised Vision Backbones with Language for Open-Vocabulary Segmentation

ICCV 2025

Honolulu, Hawaii

L. Barsellotti*, L. Bianchi*, N. Messina, F. Carrara, M. Cornia, L. Baraldi, F. Falchi, R. Cucchiara
IEEE/CVF International Conference on Computer Vision (ICCV)

ReCoptic: Computer Vision for the Reconstruction of Dismembered Coptic Codices

IEEE CH 2025

Florence, Italy

L. Bianchi, A. Moreo, F. Falchi, F. Sebastiani, C. Bianchi
IEEE International Conference on Cyber Humanities (IEEE CH) (Best Paper Award)

The Devil is in the Fine-Grained Details: Evaluating Open-Vocabulary Object Detectors for Fine-Grained Understanding

CVPR 2024

Seattle, Washington

L. Bianchi, F. Carrara, N. Messina, C. Gennaro, F. Falchi
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (Highlight Paper)

- Is CLIP the main roadblock for fine-grained open-world perception?
L. Bianchi, F. Carrara, N. Messina, F. Falchi
International Conference on Content-Based Multimedia Indexing (CBMI) (Best Paper Award)
- One Patch to Caption Them All: A Unified Zero-Shot Captioning Framework
L. Bianchi*, G. Pacini*, F. Carrara, N. Messina, G. Amato, F. Falchi
arXiv

CBMI 2024
Reykjavík, Iceland

Skills

Programming	Python, C/C++, Java, MATLAB
Deep Learning	Transformers, CNNs, Diffusion Models
DevOps & Environment	Docker, Conda, Slurm, Kubernetes, Spark, Hadoop, Unix/Linux systems, Git
Web Development	HTML, CSS, JavaScript, PHP, Flask, Apache
Libraries & Frameworks	PyTorch, Pytorch Lightning, OpenCV, scikit-learn, NumPy, Pandas, Hugging Face, Ollama, Pydantic
Databases	MySQL, MongoDB, Neo4j
Languages	Italian (mother tongue), English (fluent)

Awards

2025	Best Paper Award , ReCoptic: Computer Vision for the Reconstruction of Dismembered Coptic Codices	<i>IEEE CH</i>
2024	Best Paper Award , Is CLIP the main roadblock for fine-grained open-world perception?	<i>CBMI</i>
2024	Highlight Paper , The Devil is in the Fine-Grained Details: Evaluating Open-Vocabulary Object Detectors for Fine-Grained Understanding	<i>CVPR</i>
2015	Team Winner , ZeroRobotics Italian High School Tournament 2015 (Teamleader of TeamOre)	<i>Turin, Italy</i>

Activities

SUMMER SCHOOLS

International Computer Vision Summer School	<i>Sicily</i>
ATTENDEE	<i>July 2024</i>
• Presented a poster for the paper "The devil is in the fine-grained details: Evaluating open-vocabulary object detectors for fine-grained understanding"	

ELLIS Summer School on Large-Scale AI for Research and Industry	<i>Modena</i>
ATTENDEE	<i>September 2023</i>
• Developed a project on NeRF, evaluating its robustness in real-world scenarios.	

PEER REVIEWING

Conferences	<i>2023 - Present</i>
ICMR (2023, 2025), BMVC (2025)	

HACKATON

CTF Competitions	<i>2021-2022</i>
Member of a CTF Time ranked team (Ypsilon team).	

OTHER REPOSITORIES

ViT Patch Probing	<i>2025</i>
A linear probing pipeline to evaluate the semantic richness of ViT-based backbones (such as CLIP, MAE, DINO, etc.)	
Brain Tumor Classification	<i>2022</i>
Brain tumor detection in MRI using CNNs, ViTs, and ensemble models optimized via genetic algorithms	
Cyber Bullying Classification	<i>2022</i>
BERT-based detection for cyberbullying in tweets	
PageRank Implementation	<i>2021</i>
High-performance PageRank implementation using Spark and Hadoop, two of the most used analytics engines for large-scale data processing	
UniMusic	<i>2021</i>
Scalable music discovery platform with hybrid MongoDB/Neo4j architecture. Handles millions of tracks with real-time recommendation algorithms	
Smart Fruit Fridge	<i>2021</i>
Automation system for managing a smart fruit fridge — deployed with IoT networks using MQTT and CoAP protocols	