

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](https://github.com/lorebianchi98) | [lorebianchi98](https://www.linkedin.com/company/lorebianchi98) | [lorenzo-bianchi-893bb225a](https://www.linkedin.com/company/lorenzo-bianchi-893bb225a) | [Lorenzo Bianchi](https://www.linkedin.com/company/lorenzo-bianchi-893bb225a)

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

UNIVERSITY OF PISA

Pisa, Italy

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

UNIVERSITY OF PISA

Pisa, Italy

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

UNIVERSITY OF PISA

Pisa, Italy

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

CNR (NATIONAL RESEARCH COUNCIL OF ITALY)

Pisa, 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

DISNEY RESEARCH | STUDIOS

Zürich, Switzerland

May 2025 – Aug. 2025

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

Machine Learning Engineer

WEVO S.R.L.

Pisa, Italy

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
G. Pacini*, **L. Bianchi***, L. Ciampi, N. Messina, G. Amato, F. Falchi
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)
WACV 2026
Tucson, Arizona
- Talking to DINO: Bridging Self-Supervised Vision Backbones with Language for Open-Vocabulary Segmentation
L. Barsellotti*, **L. Bianchi***, N. Messina, F. Carrara, M. Cornia, L. Baraldi, F. Falchi, R. Cucchiara
IEEE/CVF International Conference on Computer Vision (ICCV)
ICCV 2025
Honolulu, Hawaii
- ReCoptic: Computer Vision for the Reconstruction of Dismembered Coptic Codices
L. Bianchi, A. Moreo, F. Falchi, F. Sebastiani, C. Bianchi
IEEE International Conference on Cyber Humanities (IEEE CH) (**Best Paper Award**)
IEEE CH 2025
Florence, Italy
- The Devil is in the Fine-Grained Details: Evaluating Open-Vocabulary Object Detectors for Fine-Grained Understanding
L. Bianchi, F. Carrara, N. Messina, C. Gennaro, F. Falchi
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (**Highlight Paper**)
CVPR 2024
Seattle, Washington

- 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

arXiv 2025

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

Activities

SUMMER SCHOOLS

International Computer Vision Summer School	Sicily
ATTENDEE	July 2024
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	Modena
ATTENDEE	September 2023
Developed a project on NeRF, evaluating its robustness in real-world scenarios.	

PEER REVIEWING

Conferences	2023 - Present
ICMR (2023, 2025), BMVC (2025)	

HACKATON

CTF Competitions	2021-2022
Member of a CTF Time ranked team (Ypsilon team).	

OTHER REPOSITORIES

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