Enrico Pallotta

pallottenrico@gmail.com | PallottaEnrico | enrico-pallotta | pallottaenrico.github.io | Google Scholar

Education

University of Bonn Mar 2024 – Ongoing
Ph.D. in Computer Vision *Bonn, Germany*

– Supervisor: Professor Dr. Juergen Gall.
– Topics: Video Diffusion models, Controllable Video Generation, Forecasting, World Models.

University of Bologna Sep 2021 - Oct 2023
MSc in Artificial Intelligence *Bologna, Italy*

– Core courses: *Image Processing and Computer Vision, Machine Learning for Computer Vision, Natural Language Processing, Deep Learning, Combinatorial Decision Making and Optimization.*
– Final grade: 110/110 summa cum laude.

University of Bari Sep 2018 - Jul 2021
BSc in Computer Science *Bari, Italy*

– Thesis: *"Segmentation of lung nodules with Mask R-CNN".*
– Final grade: 110/110 summa cum laude.

Work Experience

Technology Innovation Istitute Feb 2023 - Sep 2023
Research intern *Abu Dhabi, UAE*

– Perception and computer vision research intern in the autonomous robotics research center for the development of the perception module for racing drones, but also trajectory generation for control, sensor fusion algorithm, deep learning model optimization through CUDA and TensorRT.
– Working on my master thesis and a RA-L paper concerning the collection and annotation of an extended dataset for drone racing and the implementation of both deep learning and computer vision techniques for objects pose estimation.

Publications

1. **E. Pallotta**, Sina Mokhtarzadeh Azar, Shuai Li, Olga Zatsarynna, Jürgen Gall, "SyncVP: Joint Diffusion for Synchronous Multi-Modal Video Prediction", IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**) 2025.
2. M. Bosello, D. Aguiari, Y. Keuter, **E. Pallotta**, S. Kiade, G. Caminati, F. Pinzarrone, J. Halepota, J. Panerati, G. Pau, "Race Against the Machine: a Fully-annotated, Open-design Dataset of Autonomous and Piloted High-speed Flight", IEEE Robotics and Automation Letters (**RA-L**) 2024.
3. *Yuri Noviello, ***Enrico Pallotta**, *Flavio Pinzarrone, *Giuseppe Tanzi, "TeamUnibo at SemEval-2023 Task 6: A transform based approach to Rhetorical Roles prediction and NER in Legal Texts" , Proceedings of the 17th International Workshop on Semantic Evaluation (**SemEval**) 2023.

Skills

Research:	Video Forecasting, Multi-modal Video Generation, Vision-Language Models, World Models, Visual Foundation Models.
Programming:	Python (Advanced); C/C++ (Medium); Java, javascript, PHP, HTML (basic).
Tools:	PyTorch, Torch Parallel (DDP, FSDP), Linux, ROS1, ROS2, TensorRT.
Languages:	English (full proficiency); Italian (Native); German (A1).

Awards and Grants

1. **Compute award (Jupiter Supercomputer)**: 750k GH200 GPU hours for Holistic Multi-modal Egocentric Video Forecasting: working on egocentric controllable video generation for empowering Embodied AI.

Academic activities

- **Conference/Journal Reviewer**: CVPR.
- **Teaching assistant**: Video Analytics, Computer Vision, Computer Vision Lab and Seminar.