Federico Girella, Ph.D. student

☑ federico.girella@univr.it federicogirella.github.io

♥ @FedericoGirella Google Scholar

in Linkedin



About Me

I am a third-year PhD student at the University of Verona (IT), supervised by Prof. Marco Cristani 🄀 . I research joint representations in the Image and Language multi-modal domain, working with deep neural networks such as (Large) Vision and Language Models and Text-to-Image Generative Models. My main body of work focuses on Text-to-Image Retrieval and Generation in the Fashion domain. I have also collaborated for a year with Humatics SYS-DAT, developing AI tools for the fashion industry.

Education

2022 - · · · ·	Ph.D. Computer Science, University of Verona , Artificial Intelligence. AI for Vision and Language. Generative Models.
2019 – 2022	M.Sc. Computer Science, University of Verona , Visual Computing. Thesis title: <i>Multi-Task Learning: Pose Estimation for Video Analytics</i> . Final grade: 110/110L
2016 – 2019	B.Sc. Computer Science, University of Verona , Computer Science and Engineering. Thesis title: <i>Study on Augmented Reality Functionalities</i> . Final grade: 110/110

Publications

- D. Talon*, F. Girella*, Z. Liu, M. Cristani, and Y. Wang, "Seeing the abstract: Translating the abstract language for vision language models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2025.
- 2 L. Capogrosso*, F. Girella*, F. Taioli*, et al., "Diffusion-based image generation for in-distribution data augmentation in surface defect detection," in Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2024.
- F. Cunico, S. Aldegheri, A. Avogaro, et al., "Enhancing safety and privacy in industry 4.0: The ice laboratory case study," IEEE Access, 2024.
 - F. Girella, Z. Liu, F. Fummi, F. Setti, M. Cristani, and L. Capogrosso, "Leveraging latent diffusion models for training-free in-distribution data augmentation for surface defect detection," 21st International Conference on Content-based Multimedia Indexing, 2024.
- 5 L. Capogrosso, A. Mascolini, **F. Girella**, *et al.*, "Neuro-symbolic empowered denoising diffusion probabilistic models for real-time anomaly detection in industry 4.0: Wild-and-crazy-idea paper," in 2023 Forum on Specification & Design Languages (FDL), 2023.
- F. Cunico*, M. Emporio*, F. Girella*, A. Giachetti, A. Avogaro*, and M. Cristani, "Oo-dmvmt: A deep multi-view multi-task classification framework for real-time 3d hand gesture classification and segmentation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
 - M. Emporio, A. Caputo, D. Pintani, et al., "Gesture based interaction with the hololens 2," in Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter, 2023.

F. Taioli^{*}, F. Cunico^{*}, **F. Girella^{*}**, R. Bologna^{*}, A. Farinelli, and M. Cristani, "Language-enhanced rnr-map: Querying renderable neural radiance field maps with natural language," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023.



L. Capogrosso, G. Skenderi, **F. Girella**, F. Fummi, and M. Cristani, "Toward smart doors: A position paper," in *International Conference on Pattern Recognition*, 2022.

Work Experience

Humatics SYS-DAT I collaborated with Humatics SYS-DAT, implementing Deep Learning models as tools for the fashion industry. Some examples include a sales forecasting model used to predict the sales volume of unreleased fashion items, and a specialized text-to-image retrieval system.

Technologies used: Diffusion Models, Transformers, Vision and Language Models.

SAFEPLACE Project (This project aims to boost the integration and development of innovative solutions in the field of IoT Systems for healthy and sustainable living environments. My contribution resulted in the development of Computer Vision systems for the detection of crowds and face masks in challenging environments, such as shops with difficult camera angles and outdoor scenarios. These systems were deployed in real-world locations and are in use to this day.

Technologies used: Object Detection models.

Additional Activities



Skills

Languages	Strong reading, writing, and speaking competencies for Italian and English.
Coding	Python, AI libraries (e.g., PyTorch, Numpy).
Interests	Machine Learning, Vision and Language, Generative models.

Miscellaneous Experience

Certifications

2016 **Certified Level C1 in English**. Awarded by Cambridge School of English.