Florence, Italy
Media Integration and Communication Center
(+39) 370 1222 374
a_baldrati@hotmail.it
abaldrati.github.io
Alberto Baldrati
Alberto Baldrati
ABaldrati

Alberto Baldrati

PhD Student

Education

2021 - Nov, 2024 **PhD in Artificial Intelligence & Computer Vision**, *University of Pisa / University of Florence* *Media Integration and Communication Center (MICC)*, Pisa/Florence, Italy
Topics of research: Composed Image Retrieval, Vision-Language Models, Fashion Image Generation
Advisors: Marco Bertini, Andrew D. Bagdanov

2019 - 2021 **M.Sc in Computer Science and Engineering**, *University of Florence*, Florence, Italy, Grade: 110/110 cum laude
Thesis: Deep Learning techniques for image retrieval using joint textual and visual encoders

2016 - 2019 **B.Sc in Computer Science and Engineering**, *University of Florence*, Florence, Italy, Grade: 110/110 cum laude
Thesis: Video recognition of soccer actions using recurrent neural network

Work Experience

Mar, 2024 - Sep, 2024 **Computer Vision Research Scientist Intern**, *Huawei Finland Research Center*, Helsinki, Finland
Conducted a six-month internship working on video generation.

Selected Publications

2024 M. Mistretta*, A. Baldrati*, M. Bertini, A. Bagdanov. Improving Zero-shot Generalization of Learned Prompts via Unsupervised Knowledge Distillation. *European Conference on Computer Vision* ***(ECCV)***.

2024 L. Agnolucci*, A. Baldrati*, M. Bertini, A. Del Bimbo. iSEARLE: Improving Textual Inversion for Zero-Shot Composed Image Retrieval. *arXiv preprint* ***(Under review)***.

2024 A. Baldrati*, D. Morelli*, M. Cornia, M. Bertini, R. Cucchiara. Multimodal-Conditioned Latent Diffusion Models for Fashion Image Editing. *arXiv preprint* ***(Under review)***.

2023 D. Morelli*, A. Baldrati*, G. Cartella, M. Cornia, M. Bertini, R. Cucchiara. LaDI-VTON: Latent Diffusion Textual-Inversion Enhanced Virtual Try-On. *ACM International Conference on Multimedia* ***(ACM MM)***.

2023 A. Baldrati*, L. Agnolucci*, M. Bertini, A. Del Bimbo. Zero-Shot Composed Image Retrieval with Textual Inversion. *International Conference on Computer Vision* ***(ICCV)***.

2023 A. Baldrati*, D. Morelli*, G. Cartella, M. Cornia, M. Bertini, R. Cucchiara. Multimodal Garment Designer: Human-Centric Latent Diffusion Models for Fashion Image Editing. *International Conference on Computer Vision* ***(ICCV)***.

2023 A. Baldrati, M. Bertini, T. Uricchio, A. Del Bimbo. Conditioned and composed image retrieval combining and partially fine-tuning clip-based features. *Conference on Computer Vision and Pattern Recognition Workshop* ***(CVPRW)***.

2023 A. Baldrati, M. Bertini, T. Uricchio, A. Del Bimbo. Effective conditioned and composed image retrieval combining clip-based features. *Conference on Computer Vision and Pattern Recognition Demo* ***(CVPR Demo)***.

A comprehensive and up-to-date publication list is accessible on my Google Scholar profile.

Projects

2021 ***MT-BERT*** **- Multi-Task Deep Neural Networks for Natural Language Understanding**
Implement a neural language understanding model employing a Multi-Task Deep Neural Network. The model learns representations across multiple NLU tasks by utilizing BERT, a bi-directional transformer language model, as the base encoder.

2021 ***BrainAugment*****: Enhancing EEG Data through Data Augmentation Techniques**
Conducted an in-depth investigation into advanced data augmentation methodologies tailored for convolutional neural network (CNN) classification models, specifically applied to the domain of EEG data.

2019 **Video Recognition of Soccer Actions using Recurrent Neural Networks**
Python and Keras-based Bachelor's thesis project focused on training a neural network to accurately identify key soccer actions such as goals, substitutions, and cards within comprehensive match videos.

Most of the projects (in addition to others not mentioned here) are available on my Github profile.

Technical skills

Frameworks and Tools: PyTorch, NumPy, OpenCV, scikit-learn, pandas, Matplotlib

Programming Languages: Python, Java, C, C++

Miscellaneous: Git, Docker, bash

Languages

Italian, Native language

English, Full professional proficiency

Achievements

2022 **CVPR 2022 Best Demo Honorable Mention**, *Computer Vision and Pattern Recognition (CVPR) Conference*
We received a CVPR 2022 Best Demo Honorable Mention with our work "Effective conditioned and composed image retrieval combining CLIP-based features".

Other

2016 **First level instructor of Tennis, Padel and Beach Tennis**, *FIT (Federazione Italiana Tennis)*
Training course for the first-level Tennis/Paddle/Beach Tennis instructor. This course included a classroom training of 48 hours, an internship to be carried out at a club associated with the FIT, and finally a theoretical and practical examination.

2022 **Italian state professional information engineer examination**, *University of Florence*
Professional examination prescribed by Italian laws to become a professional engineer

(Last updated: October 2024)