

# Umberto Cappellazzo

Website:  
[umbertocappellazzo.github.io/](https://umbertocappellazzo.github.io/)  
Email:  
[umbertocappellazzo@gmail.com](mailto:umbertocappellazzo@gmail.com)  
LinkedIn:  
[umberto-cappellazzo-116093150](https://www.linkedin.com/in/umberto-cappellazzo-116093150)  
GitHub:  
[github.com/umbertocappellazzo](https://github.com/umbertocappellazzo)  
X (formerly Twitter):  
[twitter.com/Umberto\\_Senpai](https://twitter.com/Umberto_Senpai)

## EDUCATION

---

### University of Trento

Trento, Italy

Ph.D. in Information Engineering and Computer Science

Nov. 2021–ongoing

- Advisors: Daniele Falavigna, Alessio Brutti
- Research interests: continual learning for audio and speech processing; multi-modal (i.e., audio-language) continual learning; parameter-efficient transfer learning of audio/speech (e.g., Adapters, Mixture of Adapters, LoRA); Multi-modal LLMs for audio-visual speech recognition.

### University of Padua

Padua, Italy

MSc in Telecommunication Engineering

2016–2019

- Advisors: Michele Rossi, Matteo Gadaleta
- Thesis Title: A Deep Learning-Based ECG Delineator: Evaluation and Comparison on Standard Databases

### University of Padua

Padua, Italy

BSc in Information Engineering

2013–2016

- Advisor: Nicola Laurenti
- Thesis Title: Message Authentication over an Ideal or Noisy Channel

## WORK EXPERIENCE

---

### Jelinek Summer Workshop on Speech and Language Technology (JSALT)

Le Mans, France

Junior researcher in the FST group

June 2023 –August 2023

- Junior researcher for the “*Finite state methods with modern neural Architectures for speech applications and beyond*” group at JSALT2023 in Le Mans, France. I worked on the integration of early-exit techniques to make the training and inference of CTC/MMI systems dynamical. Our group included people from Google, JHU, Telecom Paris to name a few. More information available [here](#).

### Imperial College London

London, UK

Research Intern, *Audio-visual speech recognition meets LLMs*

February 2024 –September 2024

Supervisor: Stavros Petridis (ICL/Meta AI)

- The internship revolves around the efficient integration of LLMs for audio-visual speech recognition.

## SKILLS

---

- **Programming Languages:** Python (advanced), Java (basic), HTML (basic), Matlab (basic)
- **ML/DL Toolkits/Libraries:** PyTorch (advanced), Pytorch Lightning, NumPy, HF Transformers, Matplotlib. Good experience with CL libraries like Continuum and Avalanche
- **ASR Frameworks:** good experience with SpeechBrain and K2/icefall
- **Other:** Git, Docker

## LANGUAGES

---

- **Italian:** mother tongue
- **English:** C1
- **TOEFL:** 100/120

## MENTORSHIP & PROFESSIONAL SERVICES

---

- **Reviewer:** ICASSP 2024 [Workshop XAI-SA](#), ANNPR 2024
- **Co-supervision:** I co-supervised a MSc student from the University of Bologna (thesis title: “*On the use of Prompting for Fine-Tuning Neural Models for Speech Processing*”)

## PUBLICATIONS

---

- [1] U. **Cappellazzo**, D. Falavigna, and A. Brutti, “Efficient Fine-tuning of Audio Spectrogram Transformers via Soft Mixture of Adapters”, *Interspeech*, 2024.
- [2] U. **Cappellazzo**, D. Falavigna, A. Brutti, and M. Ravanelli, “Parameter-Efficient Transfer Learning of Audio Spectrogram Transformers”, *IEEE MLSP Workshop*, 2024.
- [3] U. **Cappellazzo**, E. Fini, M. Yang, D. Falavigna, A. Brutti, and B. Raj, “Continual Contrastive Spoken Language Understanding”, *ACL Findings*, 2024.
- [4] G. A. Wright, U. **Cappellazzo**, S. Zaiem, D. Raj, L. Ondel Yang, D. Falavigna, and A. Brutti, “Training dynamic models using early exits for automatic speech recognition on resource-constrained devices”, *Self-supervision in Audio, Speech and Beyond (SASB) Workshop, ICASSP*, 2024.
- [5] M. Yang, U. **Cappellazzo**, X. Li, S. Watanabe, and B. Raj, “Improving continual learning of acoustic scene classification via mutual information optimization”, *ICASSP*, 2024.
- [6] M. Yang, X. Li, U. **Cappellazzo**, S. Watanabe, and B. Raj, “Towards Unified Evaluation of Continual Learning in Spoken Language Understanding”, *Interspeech*, 2024.
- [7] U. **Cappellazzo**, D. Falavigna, and A. Brutti, “An Investigation of the Combination of Rehearsal and Knowledge Distillation in Continual Learning for Spoken Language Understanding”, *Interspeech (Poster)*, 2023.
- [8] U. **Cappellazzo**, M. Yang, D. Falavigna, and A. Brutti, “Sequence-Level Knowledge Distillation for Class-Incremental End-to-End Spoken Language Understanding”, *Interspeech (Oral)*, 2023.

See [Google Scholar](#) for my Google Scholar profile.