

Dr. Ismael Benito-Altamirano (<https://aiwell.uoc.edu/~ismaelbenito/>), from the AIWELL Lab at the eHealth Center, is looking for *Master Thesis candidates for a project to improve automated video editing for multi-camera events leveraging deep learning models*. The call is not restricted to only one application, see conditions below, the thesis must be carried out in 2025 (2024/25 “spring” or 2025/26 “autumn” semesters).

Who are we?

The AIWELL Lab (Artificial Intelligence for Human Well-being) group is a consolidated research group at the eHealth Center from Open University of Catalonia, located in the research facilities of the Poble Nou Campus. The group focuses on applying computer vision algorithms to health and human-centered problems, despite coming from a long way of research in the fields of scene-recognition and image understanding, the group now has recently grown with the addition of 3 new postdoctoral members and several predoctoral ones. AIWELL is in an interesting position to become a place of growth for young researchers.

Our main research resource is a shared cloud infrastructure hosted in our facilities with provides more than 16 GPU contained in a shared cluster, which enables us the training of modern neural network architectures, like transformers or large language models. Plus, the cluster is also equipped with several TB of data storage to handle modern multimodal (image, video, audio or text) datasets. Moreover, the group has experience in developing with best practices, such as the use of Git, Docker, etc., rendering a place to learn proper coding techniques alongside with your research.

Thesis proposal

The Thesis will focus on developing and validating deep learning models to automate the editing of multi-camera videos recorded during large-scale events, thanks to the technology provided by our partner, IZI RECORD (<https://izirecord.com/>). The research will utilize a novel multimodal dataset curated by professional editors. Unlike existing datasets, which primarily address complementary problems like scene change detection or video summarization, this dataset uniquely incorporates all information from the editing process itself. It captures both the raw synchronized audiovisual inputs and their professionally edited counterparts, offering unparalleled insights into the editing workflow.

The primary objective will be to design and train deep learning models capable of replicating professional editing decisions by leveraging this dataset. Advanced techniques such as multimodal pattern recognition and audiovisual synchronization will be explored to achieve precise and coherent editing results. The proof-of-concept system developed in this thesis will demonstrate the practical application of these models, validating their ability to reproduce high-quality editing outcomes in real-world scenarios.

This research addresses a critical gap in the field of automated video editing by providing a dataset that not only enhances model training but also sets a new standard for understanding the editing process. By focusing on algorithmic development and leveraging a dataset of unmatched depth and quality, the findings from this thesis will contribute significantly to advancing automated video editing technology for multi-camera setups in dynamic event contexts.

Who are you?

A highly-motivated deep-learning-enthusiast that is currently studying a Master degree in the field of computer science, mathematics, physics or electronics. Which would like to work in a research environment with a young team of researchers, in an on-growing ecosystem, which aim for remote work and a life-balanced research. With perspectives to enrol in a PhD program in a near future.

Desired skills:

- The Python language and the `numpy`, `matplotlib`, `scipy`, `pandas` stack is part of your daily workflow.
- You have experience with either `keras`, `tensorflow` or `pytorch` frameworks for deep-learning applications.
- You like to work on computer vision algorithms even if sometimes is not based upon deep learning solutions and open to work with zero-shot, few-shot and other learning strategies.
- You would like to learn best practices developing code, and focus your research in industrial applications, “making a mock-up is fun!”.

Conditions

The position is not limited to only one person at the time, variations of the same research line can be addressed if interest is raised. A position may include financial compensation depending on ongoing research projects. As for now, the maximum paid available positions are 2. Internships/contracts will follow UOC’s policies, with a variable dedication to be discussed upon acceptance of the Master thesis.

Application

Formal acceptance will follow the program circuit for the hosting Master program:

- for students of the Master in Computer Vision, please follow [these instructions](#);
- for those from the Master in Data Science, please follow [these ones](#); and,
- for other program, we will resolve the proper circuit to enable the direction of the Master Thesis.

In any of the scenarios, please contact Dr. Ismael Benito-Altamirano at ibenitoal@uoc.edu before applying with a CV and a brief email of interest.