



Master in  
Computer Vision  
Barcelona

# M5 Project: Cross-modal Retrieval

## Week 2

Introduction to Object detection and  
Instance Segmentation with Detectron2

Rubèn Pérez Tito  
[rperez@cvc.ua  
b.cat](mailto:rperez@cvc.ua<br/>b.cat)

Ernest Valveny  
[ernest@cvc.ua  
b.cat](mailto:ernest@cvc.ua<br/>b.cat)

# General comments on the report

- In general, good quality of writing and organization of the report
- Introduction should contain a brief description of the task (detection and segmentation), context and applications, and a summary of the work that you will explain in the report. It can contain a generic overview of existing methods, but put all details about the methods in the section on related work
- Related work should include all different types of approaches to object detection and segmentation (one-stage, two-stage, based on transformers, ...). It should not be only a list of descriptions of different methods without any relation among them. Try to include some categorization of methods and give information about relations between different methods (advantages, drawbacks in terms of application, performance, ...)

# Specific comments on the report

- Group 1: Very good introduction: historical context, applications, definition of the tasks and summary of the work on object detection and segmentation. Related work includes discussion of methods and notes on accuracy, speed, etc.
- Group 2: Move description of faster and mask r-cnn from introduction to related work. Give some structure to related work. Categorize methods and put them in relation. Add references.
- Group 3: Change format to recommended CVPR format. Include description of transformer-based methods in related work.
- Group 4: Add an introduction to the tasks on object detection and segmentation. Extend a bit more description on transformer-based methods.
- Group 5: not necessary in the report to talk about the task on image classification. Give a better structure to related work. For instance, faster and mask r-cnn are covered in two different parts of the section. You could give a more specific name to section 2.3 and create a new subsection for works on segmentation.
- Group 6: Change format to recommended CVPR format. Related work is uncomplete. Only family of r-cnn methods is included. Include Also one-stage, segmentation and transformer-based methods . Categorize methods and give structure to the related work.
- Group 7: Good introduction and good structure for related work. Good diagram showing different types of architectures. Include some description of transformer-based methods