

Exercise: YOLO – You Only Look Once (ca. 1-2h, depends on how fast you read & understand...)

## 1. Learning goal

In the lecture we have talked about three different object detection models: R-CNN, Fast R-CNN and Faster R-CNN. However, many more CNN based object detection models have been proposed in the last years. Another important model is YOLO. The goal of this exercise is to understand how it works and how it differs from the three object detection models that I presented in the lecture.

## 2. Read the YOLO paper

First read the original paper:

Redmon, J.; Divvala, S. K.; Girshick, R. B. & Farhadi, A.  
*You Only Look Once: Unified, Real-Time Object Detection*

## 3. Questions

- 1.) How does YOLO work?
- 2.) How many convolutional layers does YOLO use?  
How many layers are used for the MLP at the end?
- 3.) What are the advantages of YOLO compared to R-CNN, Fast R-CNN and Faster R-CNN?
- 4.) For what are the different parts of the “multi-part loss function”?
- 5.) What are the limitations of the YOLO model?
- 6.) How fast is YOLO? How good is YOLO compared to other models?