Fruit Detection

Lorena Lazar, Katarina Lukašević, Luko Lulić, Laura Petan, Tea Teskera

Introduction

- Detection
  - Localization of the object
  - Classification of the object
- Deep learning methods
  - YOLOv8
  - Faster R-CNN
- Traditional methods
  - support vector machine
  - k-nearest neighbours

Problem description

- Localization and classification of fruit on a dataset of natural images
- Comparison of deep learning and traditional methods

Applications of fruit detection

- in agriculture, for automatic fruit sorting
- in supermarkets, instead of barcodes
- as a tool in primary education and for working with people with Down's syndrome

Overview of existing methods

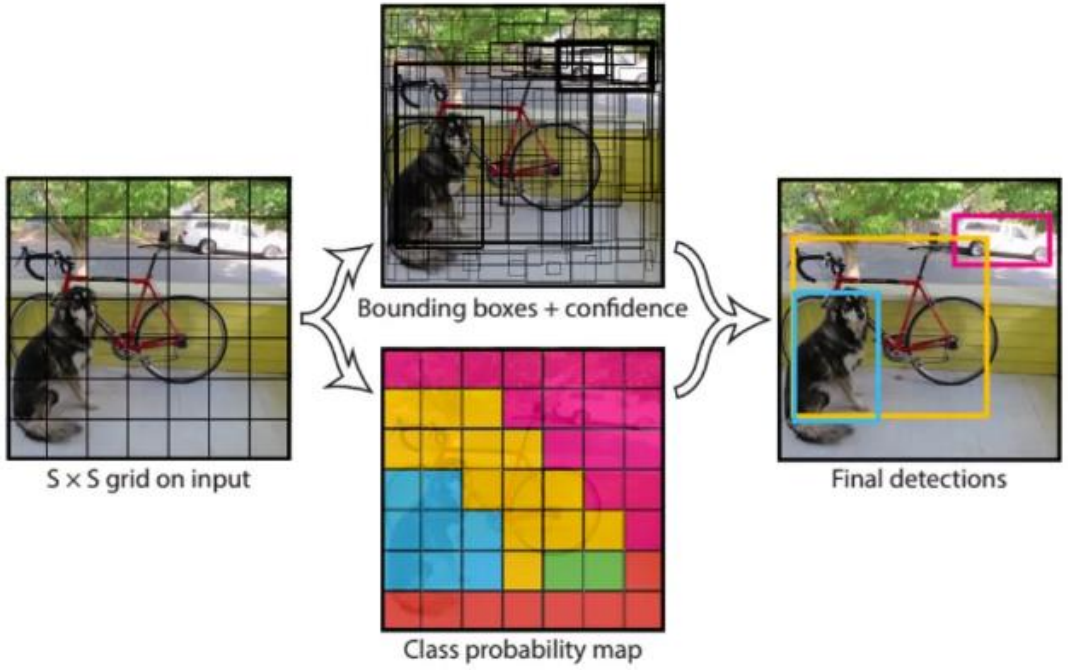
- Deep learning methods using convolutional neural networks such as YOLO and R-CNN
- Traditional methods such as Support Vector Machines (SVM) and k-Nearest Neighbours (k-NN)
- Laboratories and groups
  - Many published articles exist regarding these methods
  - Ultralytics company (created YOLOv8)

Dataset

- Publicly available dataset on Kaggle
- 300 .jpg images of fruit and 300 .xml files with bounding boxes and classes present in each image
- 3 classes of fruit: apple, banana, orange
- 240 training set – 60 testing set

Traditional ML Methods

- Support Vector Machine
  - relatively memory efficient, not suitable for larger and more complicated datasets
- K-Nearest Neighbours
  - no training period, easy implementation
  - Does not work well with larger datasets, higher dimensionality or noise in images



YOLO

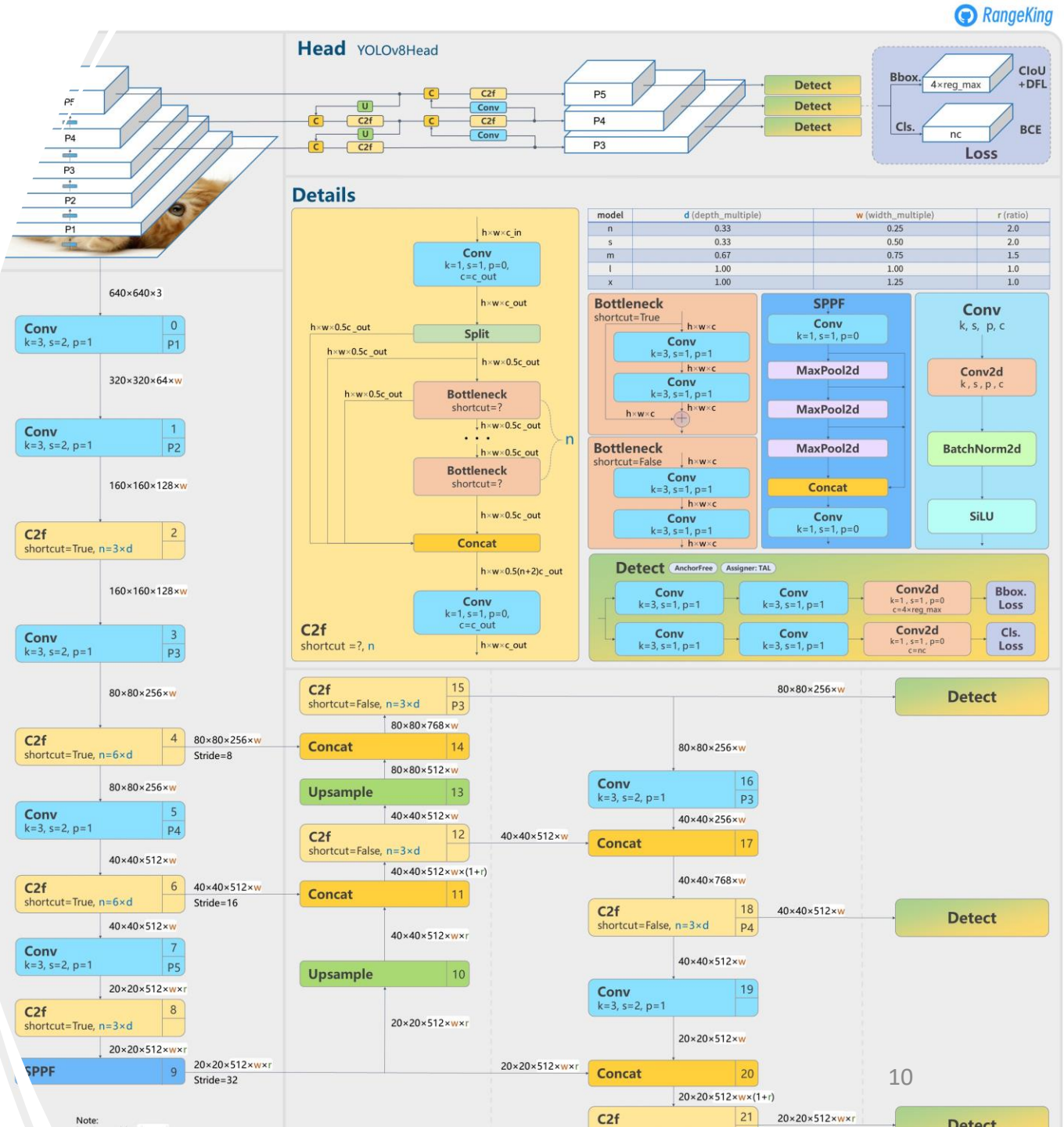
- *You Only Look Once*
- A family of deep learning algorithm which employ convolutional neural networks
- Three main methods:
  - residual blocks
  - boundary box regression
  - intersection over union

YOLO

- Advantages:
  - real-time
  - one pass through the network is enough
- Disadvantages:
  - unable to detect new or unusual shapes
  - cannot reliably detect small objects in groups

YOLOv8

- Developed by Ultralytics, released in January of 2023
- Comes in 5 sizes – biggest one has 68.2 million parameters
- Three major changes in comparison to previous iterations:
  - anchor free detection, introduction of C3 convolutions, mosaic data augmentation



R-CNN

- Constant starting number of regions
- Greedy recursive algorithm groups regions
- Uses CNN and SVM
- Fast R-CNN and Faster R-CNN

Faster R-CNN

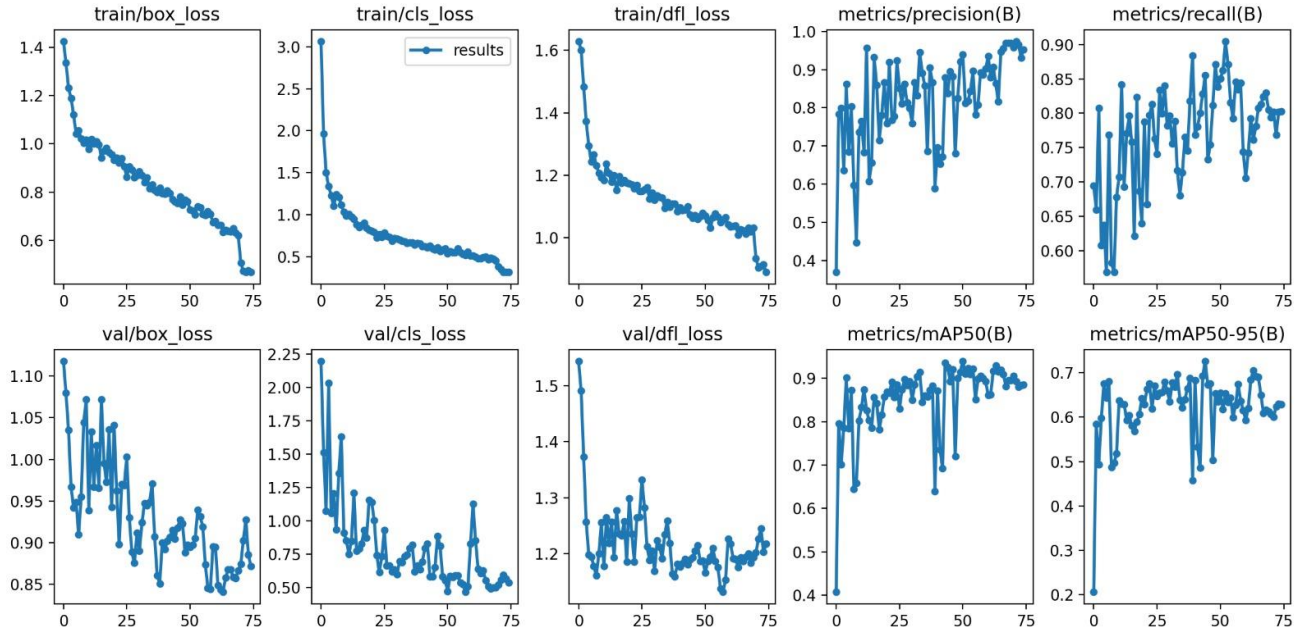
- Advantages:
  - eliminates the need for selective region look-up
- Disadvantages:
  - high data requirements – needs large amounts of annotated data to train and fine-tune

Experimental results with traditional methods

- Support Vector Machine
  - F1-value of 0.83, 0.91 and 0.85 for apples, bananas and oranges, respectively
- K-Nearest Neighbours
  - F1-value of 0.78, 0.89 and 0.82 for apples, bananas and oranges, respectively

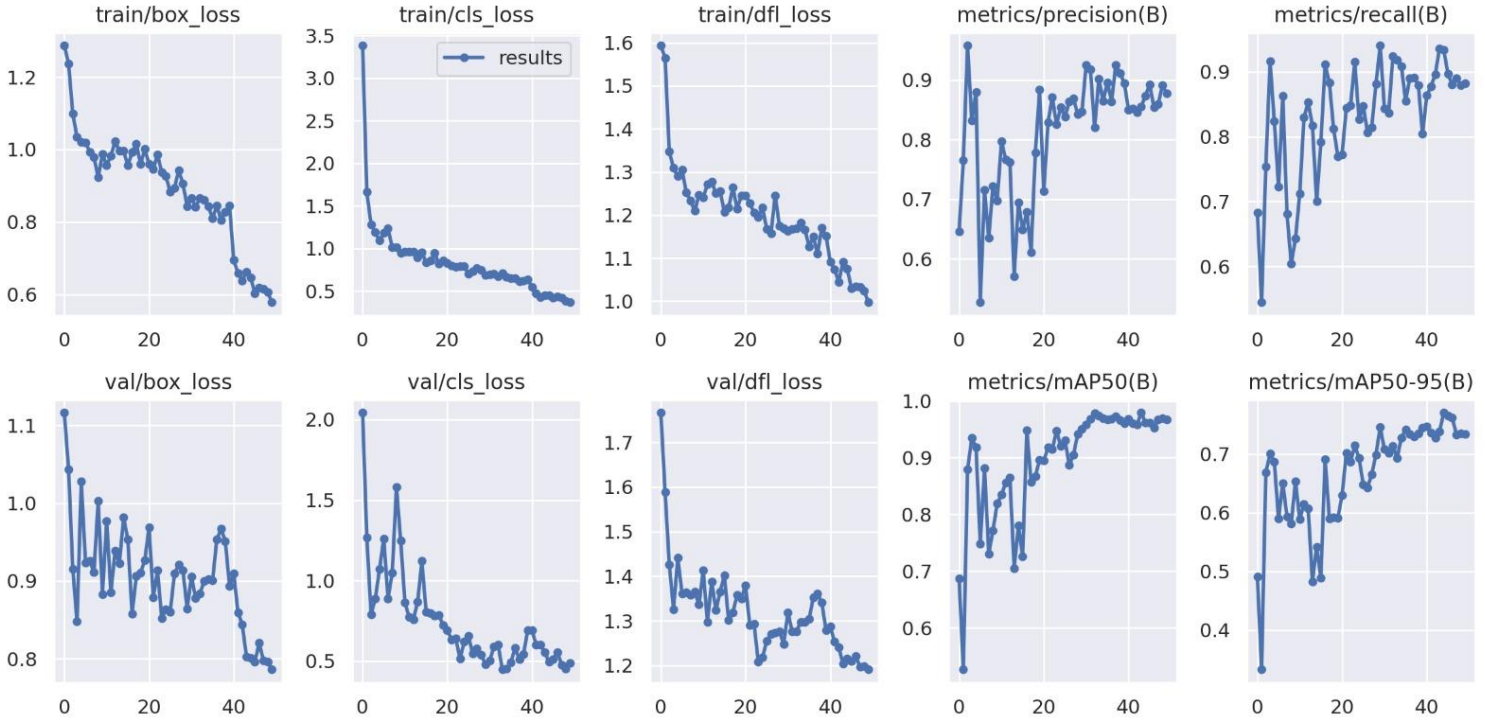
YOLOv8s results

	Precision	Recall	mAP50	mAP50-95
Total	0.837	0.855	0.926	0.728
Apple	0.668	1	0.995	0.895
Banana	0.874	0.696	0.795	0.487
Orange	0.969	0.868	0.987	0.802

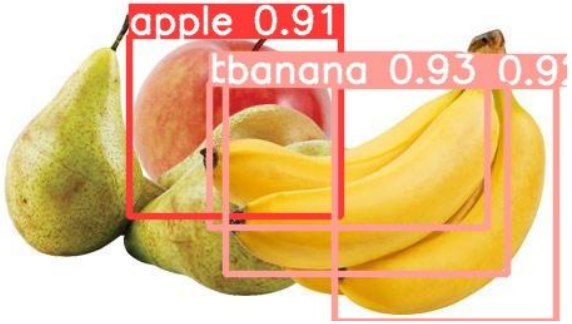


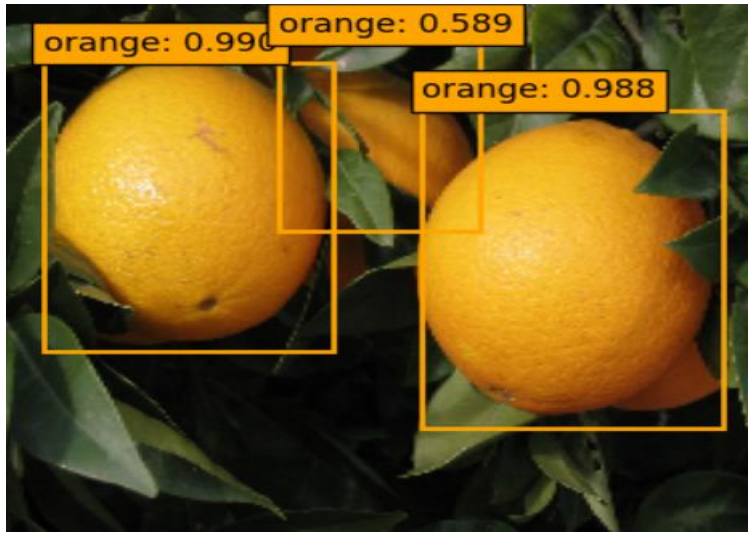
YOLOv8x results

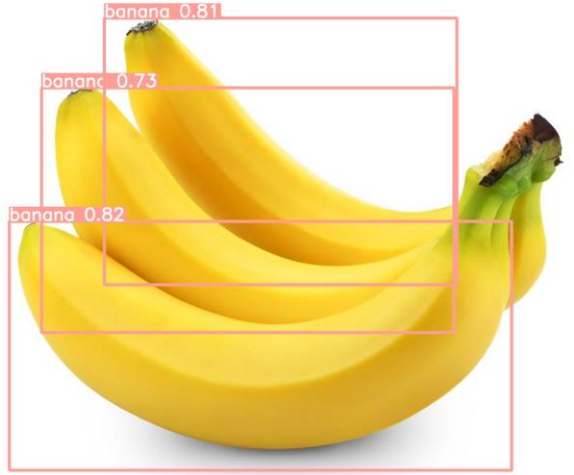
	Precision	Recall	mAP50	mAP50-95
Total	0.874	0.934	0.962	0.77
Apple	0.691	1	0.995	0.895
Banana	1	0.886	0.919	0.597
Orange	0.94	0.917	0.972	0.816

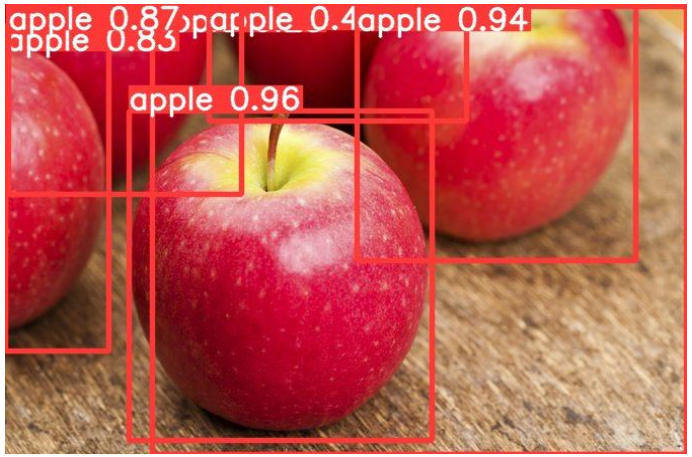


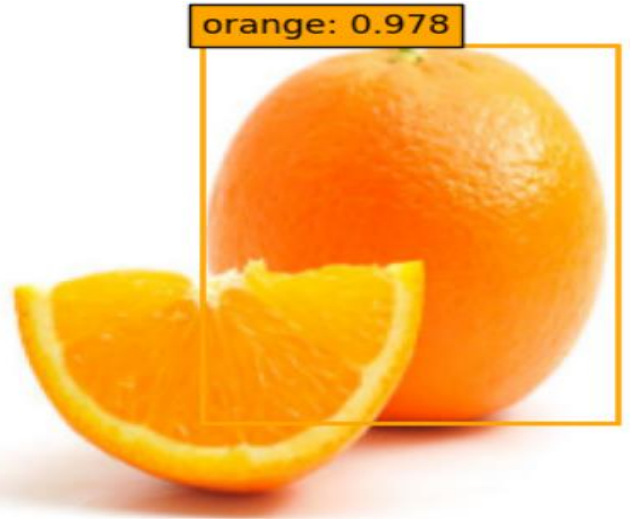
Examples of successful detection



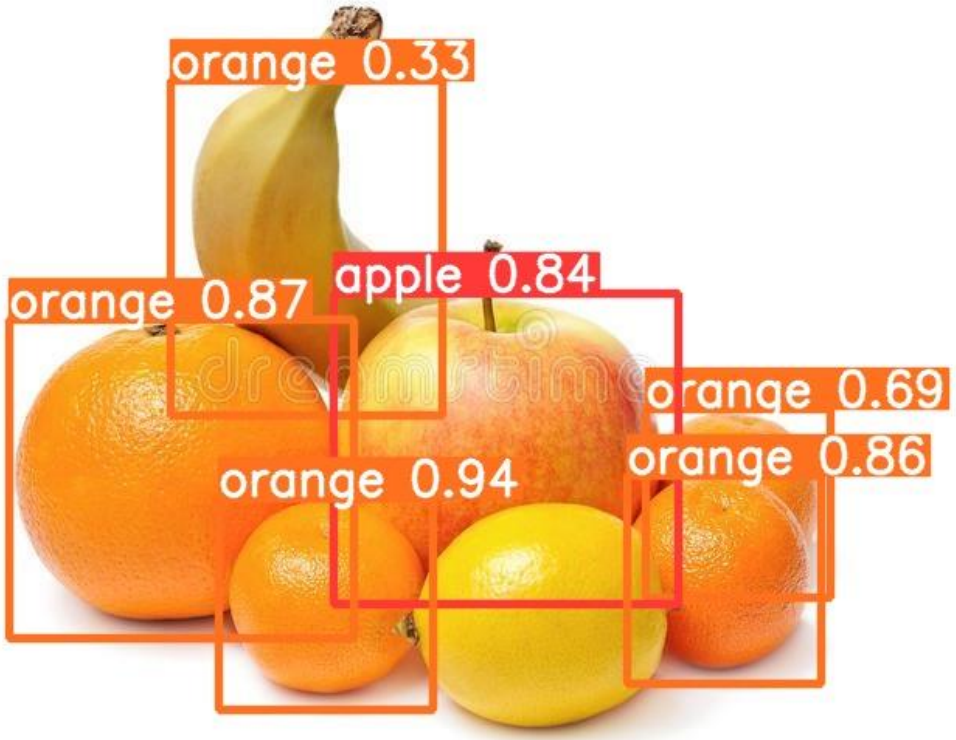


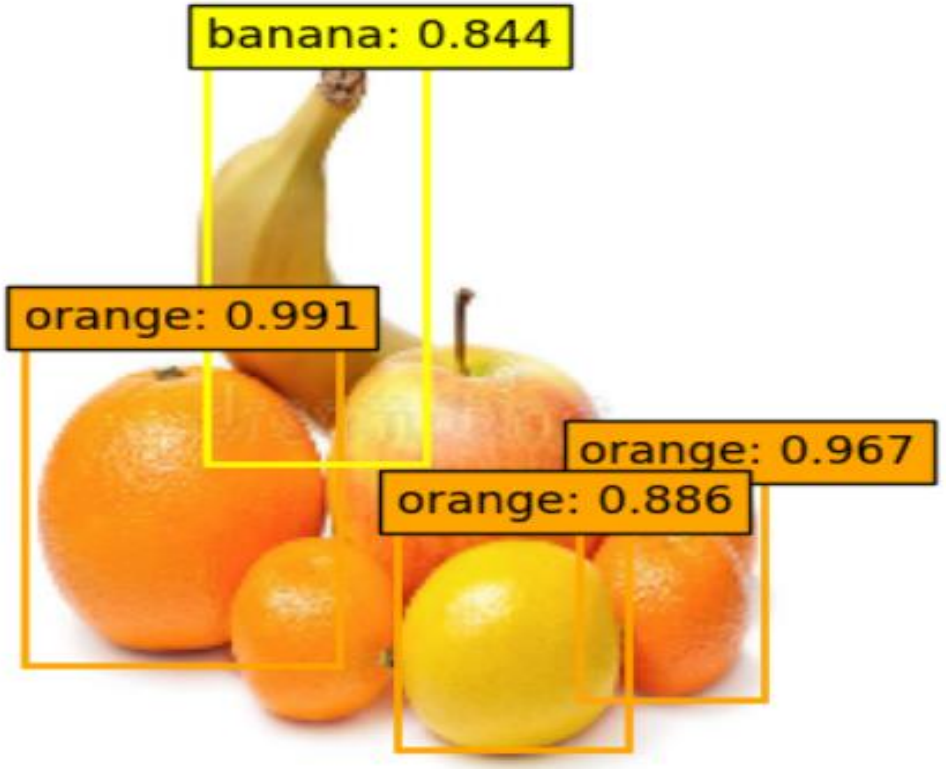




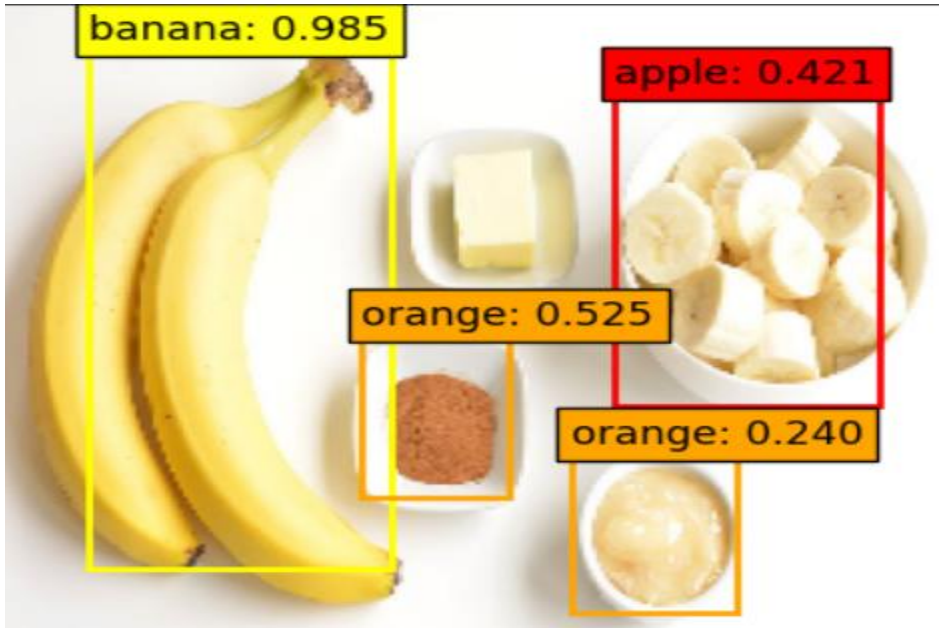


YOLOv8 vs Faster R-CNN





YOLOv8 vs Faster R-CNN



Conclusion

- YOLOv8 performs best
  - mAP50: > 96%
  - Inference time: 12.5 ms
- Future works
  - Try out different YOLO versions
  - Use a larger and more complicated dataset

Q&A time