Logo Detection Project

20600 DEEP LEARNING FOR COMPUTER VISION
Bocconi University
Gaia Rubera, Francesco Grossetti

Steffen Brockmann - 3150953
Alessandro Caruso - 3054408
Silvia Juzova - 3136278
Fabian Kraus - 3032386
Louis Lacombe - 3053198
Riccardo Tordini - 3057162

Problem definition

Task

Object detection of brand logos

Dataset

39,000 instagram images containing 17 logo classes

Evaluation

Intersection over Union (IOU)

Roadmap

Step 1

Preliminary Data Analysis

Step 2

Data Preprocessing

Step 3

Description of models deployed

Step 4

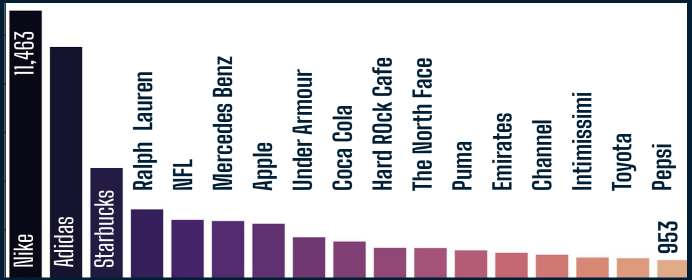
Performance evaluation

01

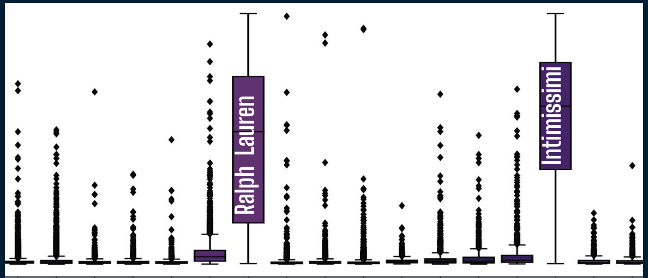
Preliminary Data Analysis

Preliminary Data Analysis

Distribution of labels



Distribution of bounding box area

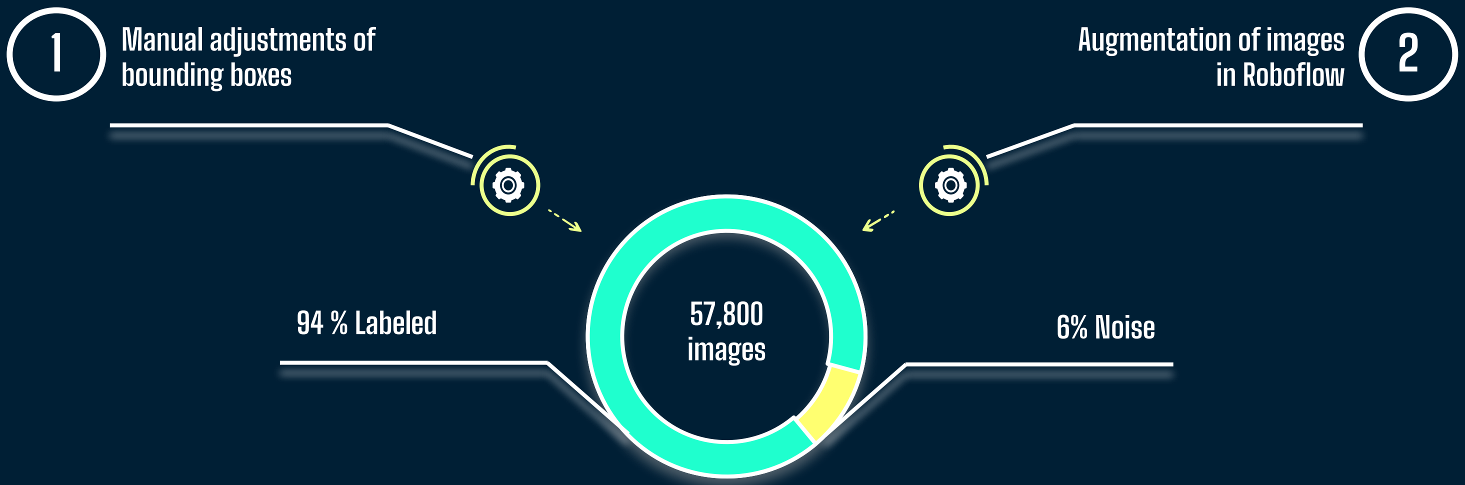


Class imbalance	**Generate additional augmented images or drop**
Wrong labels for: *Ralph Lauren & Intimissimi*	**Use as noise data**
Bounding boxes too high	**Drop from the dataset**
Images not found in the image folder	**Drop from the dataset**

02

Data Preprocessing

Preprocessing of training data



03

Models

a. Faster R-CNN
b. YOLOv5

Pre-training model comparison

	Faster R-CNN	YOLOv5
Structure	Two-stage	One-stage
Time efficiency	Low	High
Expected accuracy	Medium	High

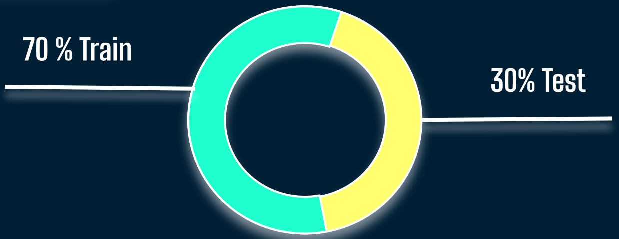
Models' specification comparison

Faster RCNN

Parameters:

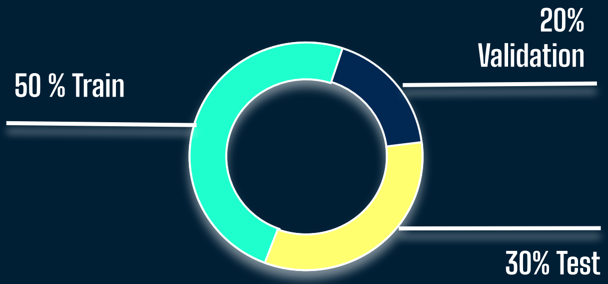
Image size:	800
Box batch size per image:	36
RPN batch size per image:	24
NMS threshold:	Lowered
Aspect ratios:	Based on previous research

Image split:



YOLOv5

Image size:	640
Learning rate:	0.01
SGD Momentum:	0.938
Weights	Nano & XLarge
Extra augmentation:	Degrees, flipud, ...

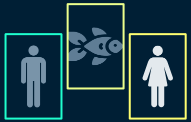


04

Performance Evaluation

Justification of the approach adopted

Multiple unlabeled logos present in the pictures



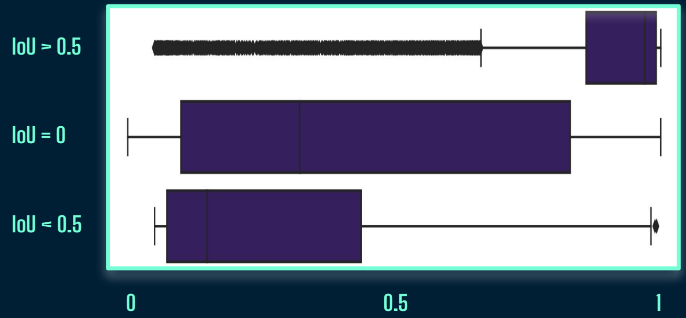
Faster RCNN & YOLOv5 are designed to predict multiple instances

Evaluation based on predictions with highest confidence would understate the predictive power

$$\text{Confidence}_{\text{BBox } i} > \lambda \Rightarrow \text{True Logo}$$

Choosing confidence threshold

Faster RCNN



Chosen threshold:

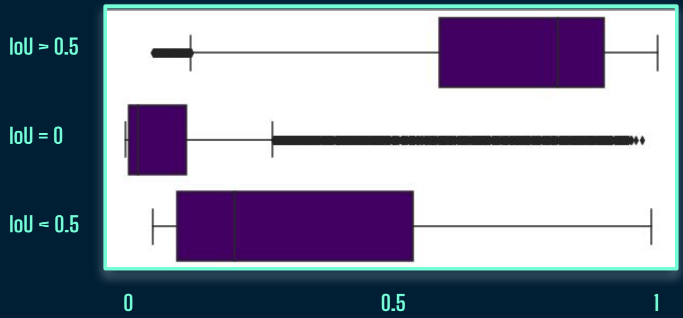
7.34%

% of incorrect predictions
~13.4%

# of predictions per picture:
1.23

Many true predictions with 0 IoU & high confidence.

YOLOv5



Chosen threshold:

8%

% of incorrect predictions
~18%

# of predictions per picture:
1.3

Producing a lot of predictions where it is uncertain.

Predictions based on highest confidence

Not representative score, as **multiple labels** in pictures

Predictions based on highest IoU

- Yolo overperforms Faster RCNN
- Logos with lowest IoU, also labels with lowest confidence
- Nike's bad performance due to its features

Examination of result differences

? Instances of other logos not labelled in the images?

Filter predictions

Predictions with **higher confidence score than** the one associated with **the highest IoU**

Create database

578, 623 predicted boxes for Faster RCNN and Yolo respectively

Manual inspection

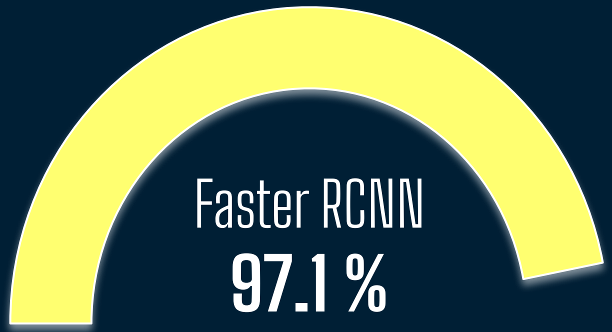
91%

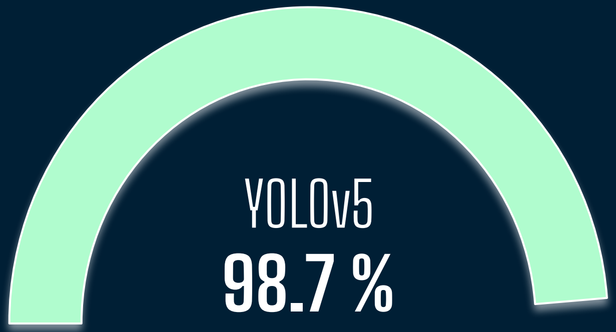
Faster RCNN correctly identified labels

89%

YOLOv5 correctly identified labels

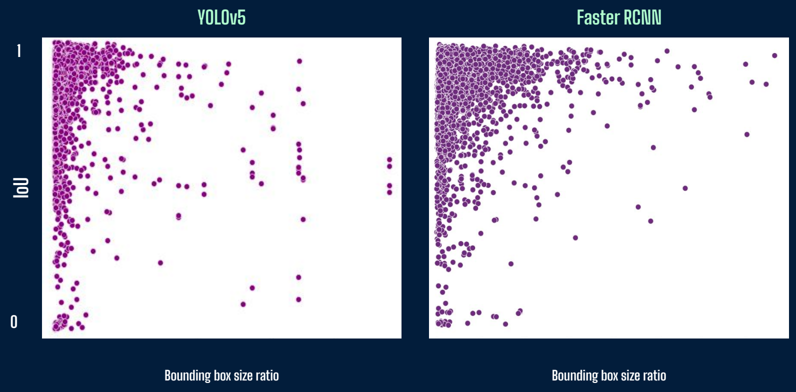
Classification accuracies achieved





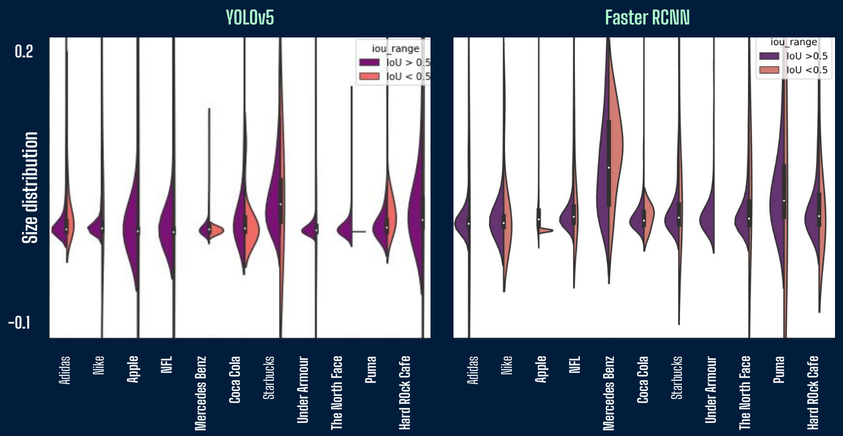
Exploring systematic errors

Correlation plot between IoU and box size



Correlation between the size of the bounding box and the prediction performance ⇨ Stronger correlation for Faster RCNN, consistent with empirical research.

Distribution plot of box sizes across different IoU



For some logos issue to predict the large bounding boxes too. ⇨ **Starbucks:** average of bounding boxes with low IoU is at 0.04 vs. 0.015 for high IoU.

Findings deduced from the analysis

FRCNN

Higher confidence in the predicted labels and tendency to **predict less false** bounding **boxes.**

YOLO

Trade-off between **predictive performance** and additional **error in predicted** bounding **boxes**

Both

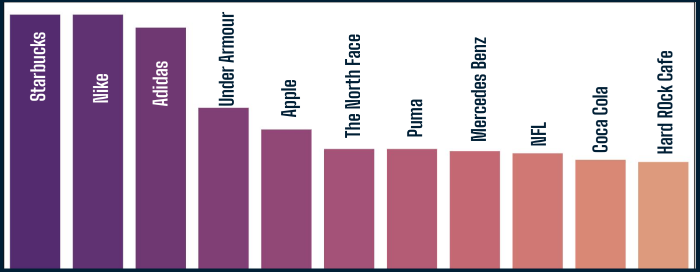
Using **highest confidence understates the results** of the model.

Both

Test on an additional testset, where all **predictions > λ%** will be considered as **valid logos.**

Designing extra test set

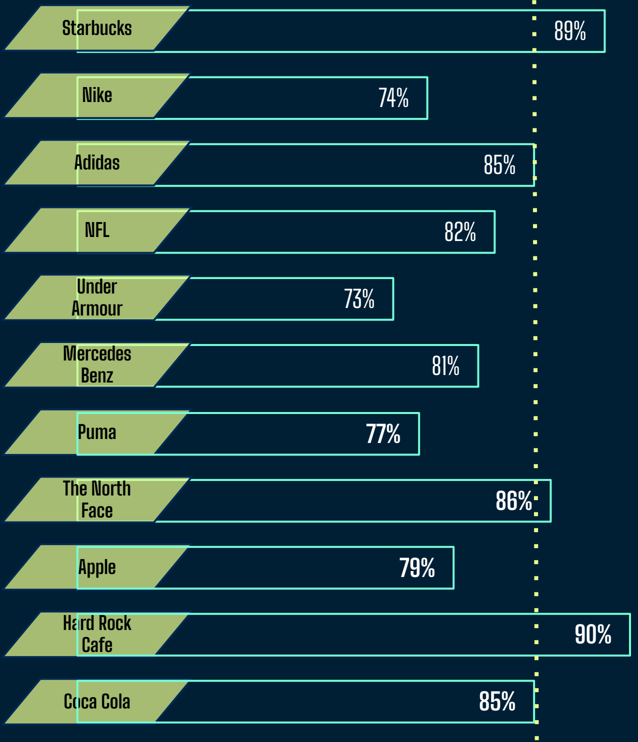
Distribution of labels in additional test set



Original test set contained many images with **no or wrong labels**, **inaccurate bounding boxes** and only **one of multiple possible labels**

Download **additional 600 images** from all classes and labeled them accurately, **majority had multiple instances of logos**

Yolov5 on multi-logo detection task



Findings deduced from the analysis

Logos with high scores

Unique design characteristics

Logos with low scores

Easy to be mistaken for basic shapes.

Pictures with too many labels

Pushing the average value to lower levels.

Incorrect Predictions: 7%

Fractions of predictions that are not logos

Conclusions & Recommendations

- **YOLOv5 outperforms** Faster R-CNN in terms of both accuracy and speed
- Less **unique shapes** (e.g. Nike) are generally **harder to detect** as well as **smaller bounding boxes**
- **Extensive preprocessing** and **improving** the dataset was essential for good results

- **Relabeling** of the dataset providing annotations and bounding boxes for **multiple logos** in the image for multi-object detection
- **Remove** ethically **questionable images,** utilize only balanced and clean dataset for real-life applications

THANKS FOR YOUR ATTENTION