Detecting Swimming Pools using Aerial Imagery

- Cruoglio Antonella - 2025992
- Iovino Giuliana - 2017512
- Mascolo Davide - 2001991
- Napoli Mario - 2015169

Data Source

Dataset available on kaggle:

The **dataset** is constituted by 1224 images:

- Original size: 25,000 x 25,000 pixels
- Patches of 512x512 pixels
- 3197 annotated pools with different shapes and hues

It is splitted into **train set**, **validation set** and **test set**:

Train	80% of the images	980 images
Validation	10% of the images	122 images
Test	10% of the images	122 images

Pipeline

- **Preprocessing steps:** extract information related to bounding boxes xml file with PASCAL VOC format.

- Experiment our models with some parameters

- **Data Augmentation:**
  1. image flipping
  2. image rotation
  3. image motion blur
  4. image median blur

Configuration model

- Device : cuda, type Tesla T4
- Batch Size: 8
- Image Size: 416
- Epochs: 10
- Stochastic Gradient Descent with:
  - learning rate = 0.001
  - momentum = 0.9
  - weight decay = 0.0005

Models

CNN based object detection algorithms can be divided into two major categories:

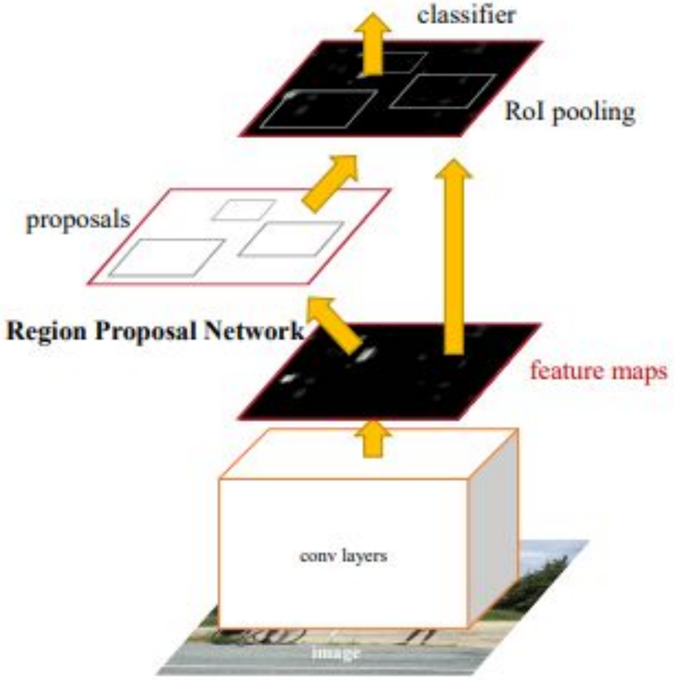
- One stage detector:
    a. **Retina-Net** (one stage)
    b. **YOLO** (one stage)

- Two stage detector:
    a. **Faster R-CNN** (two stages)

Faster R-CNN: two stage detector.

It constructs a Region Proposal Network to predict proposals from features.

- Simultaneously regress region bounds and objectness scores
- Predict multiple region proposals called anchors.
- Combines four losses

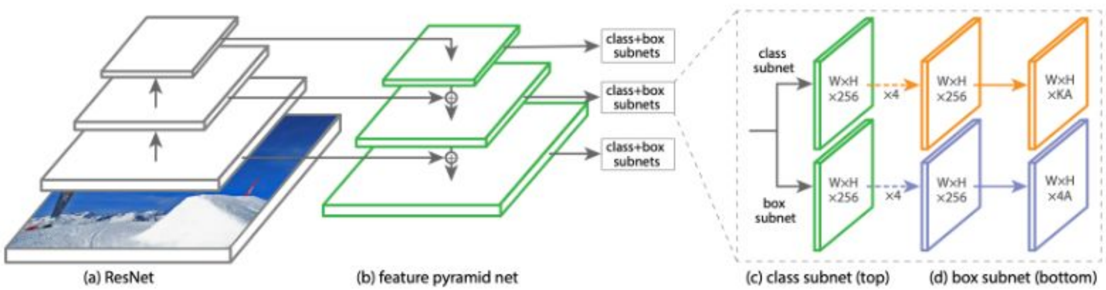


[Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks]

RetinaNet: one stage detector.

is one of the best one-stage object detection models that has proven to work well with dense and small scale objects. It is a single, unified network composed of a *backbone* network and two task-specific *subnetworks*.

- The backbone includes:
    1. The ResNet
    2. The Feature Pyramid Network
- The Classification Subnet
- The Regression Subnet.



[Focal Loss for Dense Object Detection]

YOLO: one stage detector.

It is an object detection method that requires only a single forward propagation through a neural network to detect objects.

- Detection as Single Regression Problem
- Developed as Single Convolutional Network
- Reason Globally on the Entire Image
- Learns Generalizable Representations

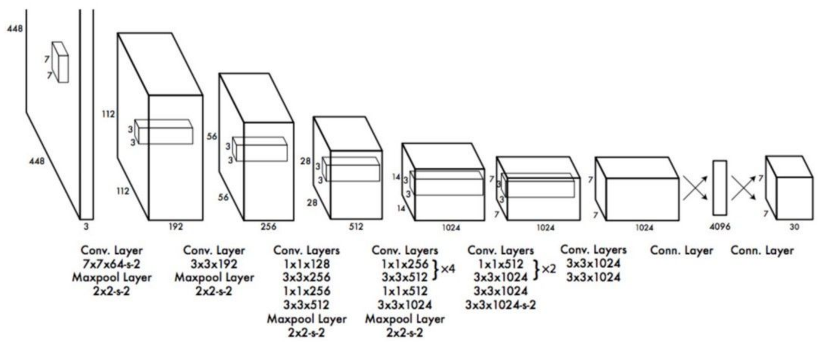


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

You Only Look Once: Unified, Real-Time Object Detection

Results

	Precision	Recall	F1-Score
Faster R-CNN	0.911	0.965	0.937
RetinaNet	0.918	0.959	0.938
YOLO	0.944	0.924	0.933

Training Time:

- Faster R-CNN: ~34 mins
- RetinaNet: ~30 mins
- YOLO: ~4 mins

Examples of good predictions with Faster R-CNN

Ground Truth

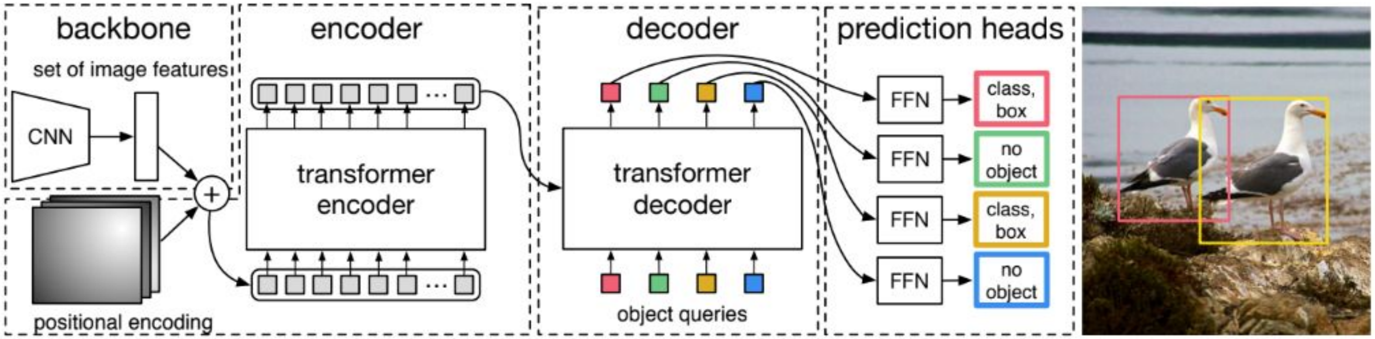
Example of FP in Faster C-NN

Example of FP in YOLO

Further works

DETR model

- Object detection as a direct set prediction problem
- Set-based global loss that forces unique predictions via bipartite matching
- Transformer encoder-decoder architecture



[End-to-End Object Detection with Transformers]

Conclusions

All the models reach good performances.

- We want to minimize false negatives maximizing **Recall**.
- We want to minimize false positives maximizing **Precision**.
- Considering that we are interested in minimize both we can also look at **F1-Score.**

References

- S.Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2015.

- J. Dai, Y. Li, K. He, and J. Sun, R-FCN: object detection via region-based fully convolutional networks, 2016.

- J. Ding, N. Xue, G. Xia, X. Bai, W. Yang, M. Y. Yang, S. J. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang. Object detection in aerial images: A large-scale benchmark and challenges, 2021.

Thank you