YOLO – a fast object detector

Alexandr Ignatenko
Hamburg, 27.01.2020

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

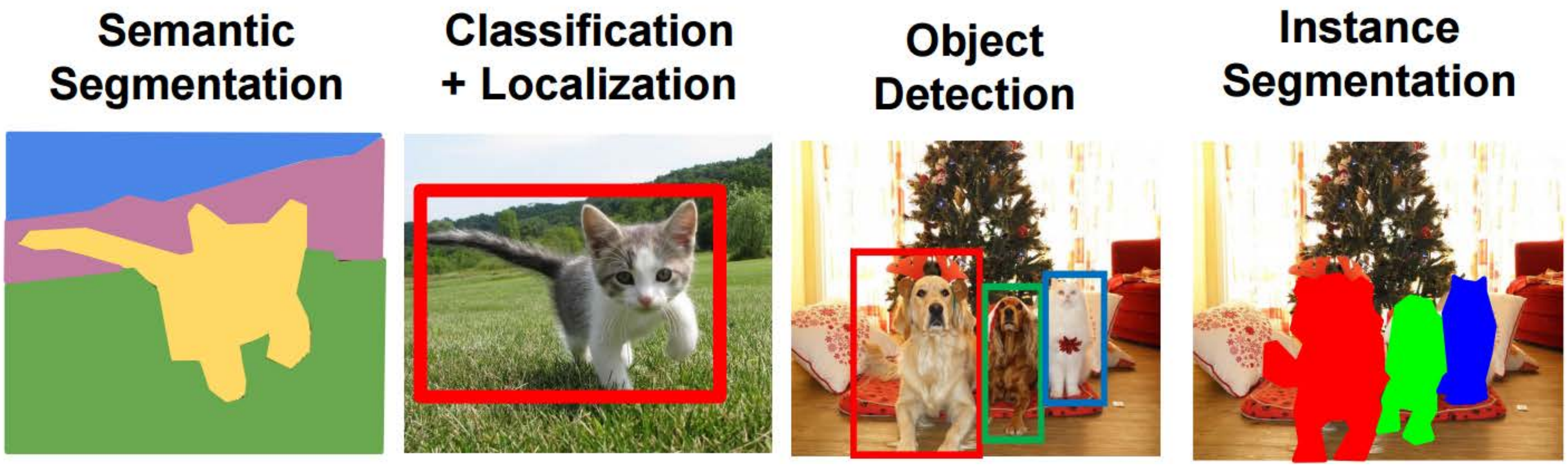
DESY.

Overview

- **Introduction**
- **Popular object detectors**
- **YOLO**
- **Application for classification of diffraction patterns in SPI experiments**
- **Conclusions**

Introduction

Object detection among computer vision tasks

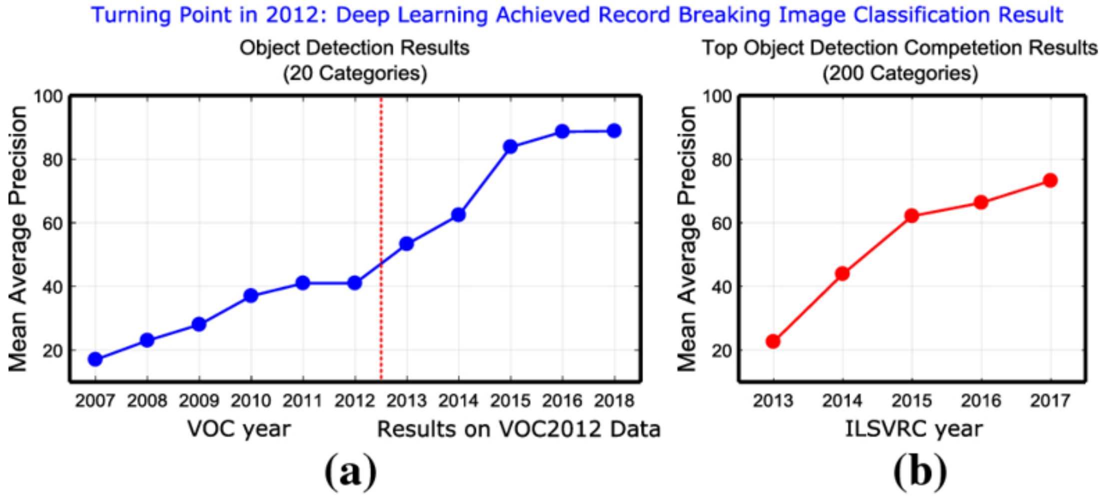


Object detection: given an input image, predict the locations of a certain class of objects in the image

Locations are usually represented using bounding boxes

Historic object detection

Out of scope



L. Liu et al., Int J Comput Vis (2019)

Object detection before 2012:

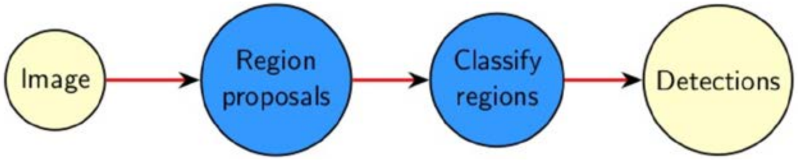
Define region proposals → Calculate features for each region → Classify each region

- Handcrafted features:
    - Haar-like
    - Histograms of Oriented Gradients (HOG)
    - …
- Region proposals
    - Sliding window
    - Selective search
    - …

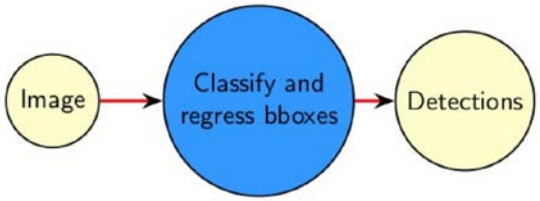
CNN era

Object detection after 2012:

- 2-stage methods
  - R-CNN
  - Fast R-CNN
  - Faster R-CNN
  - …
- Single-shot methods
  - SSD
  - YOLO
  - …

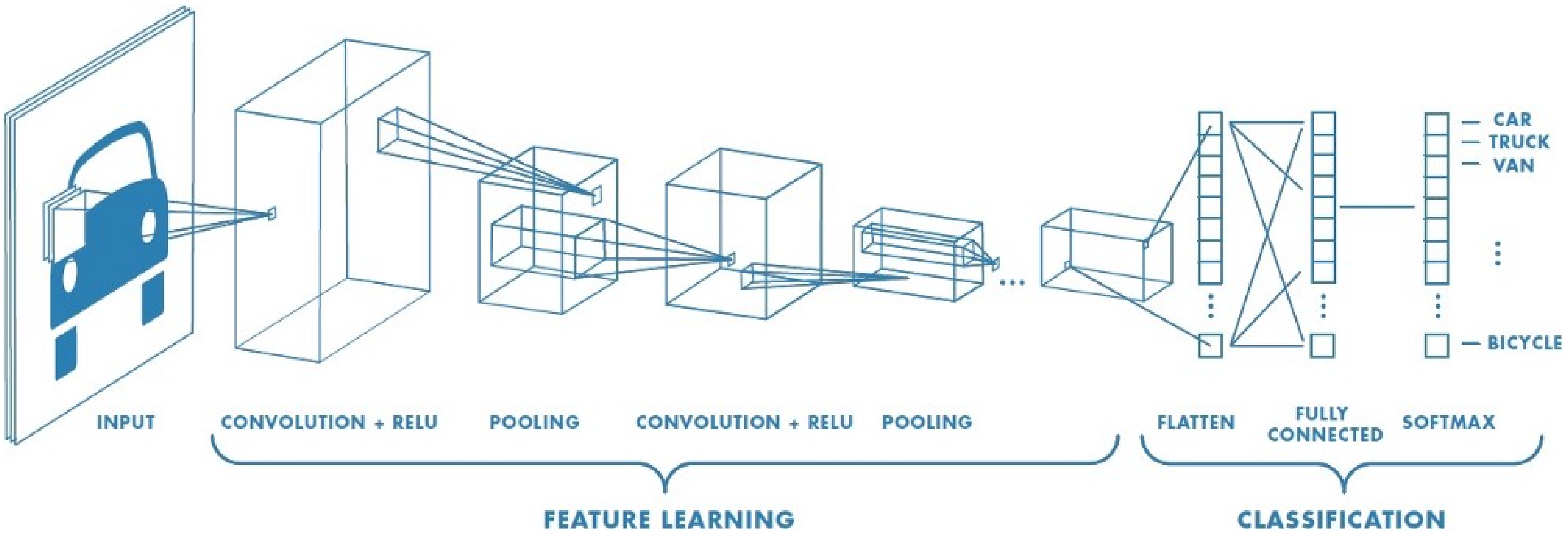


Talk of T. Kerola (2019)



Understanding of CNN

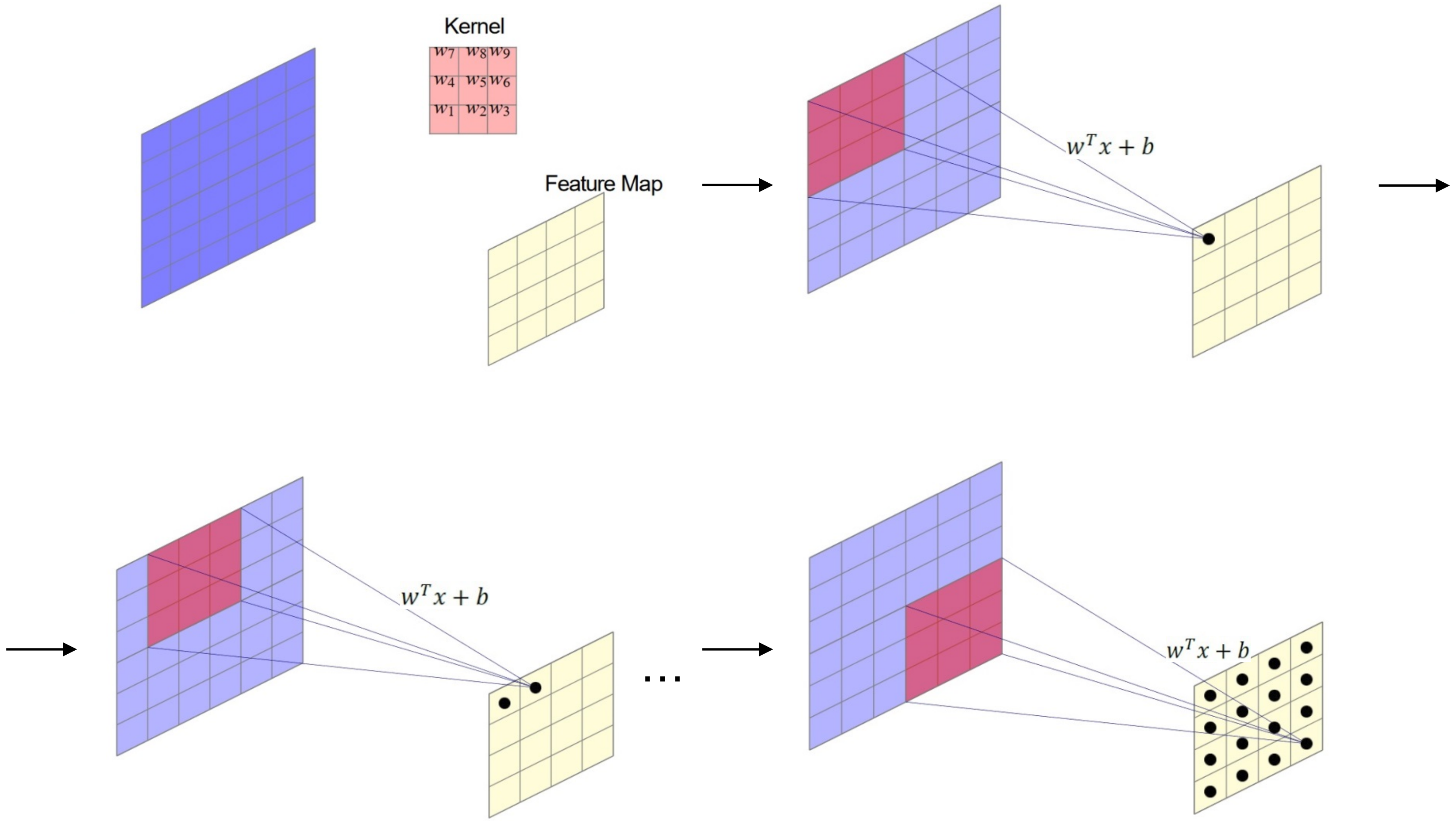
CNN for classification



- Feature learning part is trained on the largest annotated data set
- Classification part is trained for custom classes

Understanding of CNN

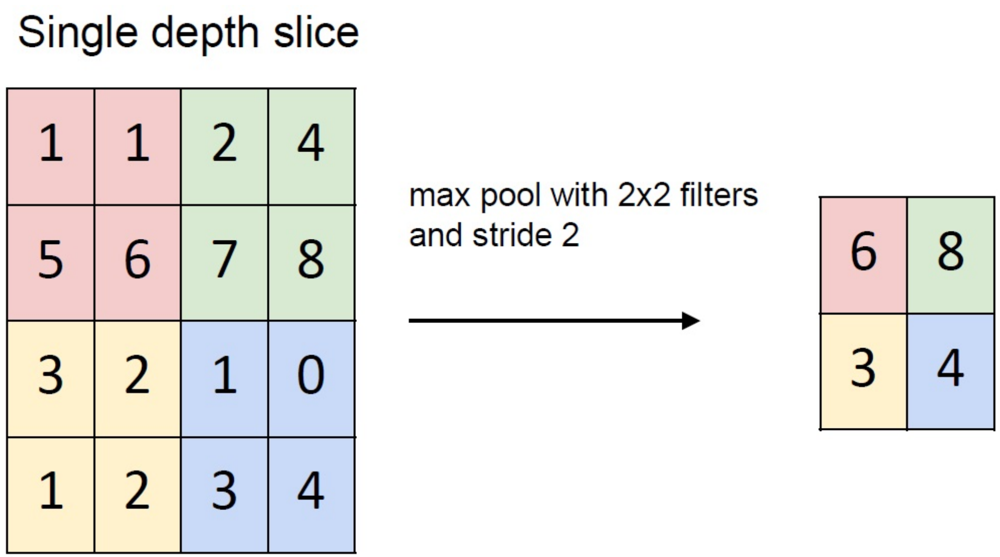
Convolutional layer



http://cs231n.stanford.edu/

Understanding of CNN

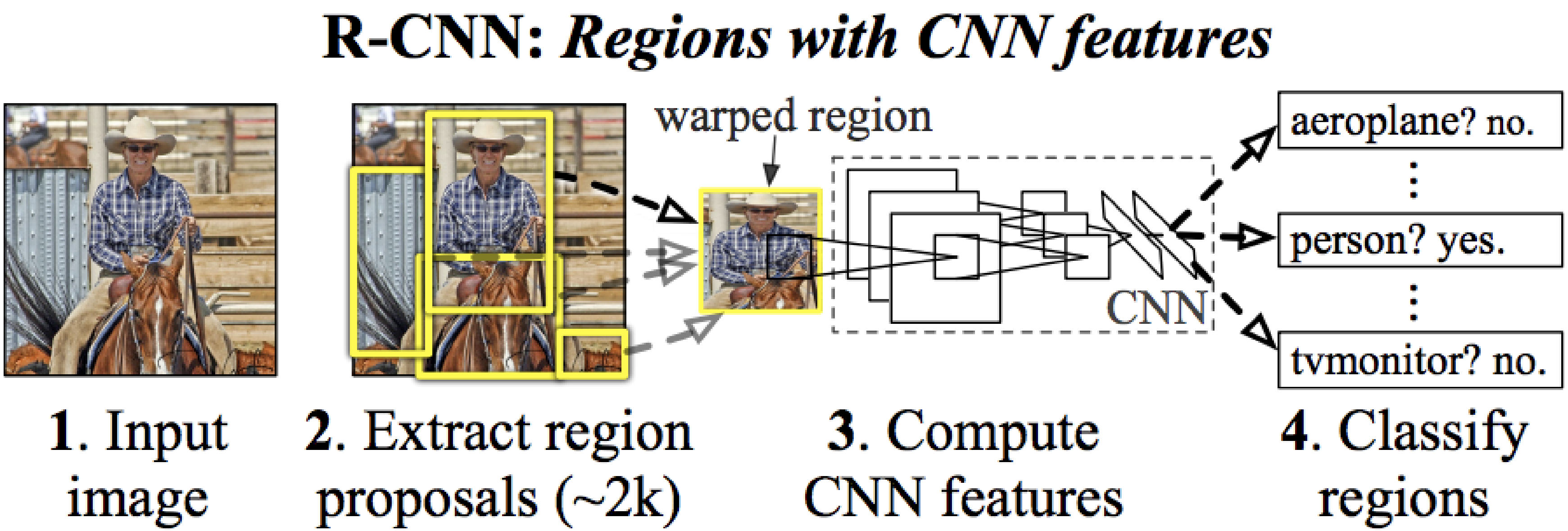
Pooling layer



Other pooling functions: min, average, L2 norm …

Popular object detectors

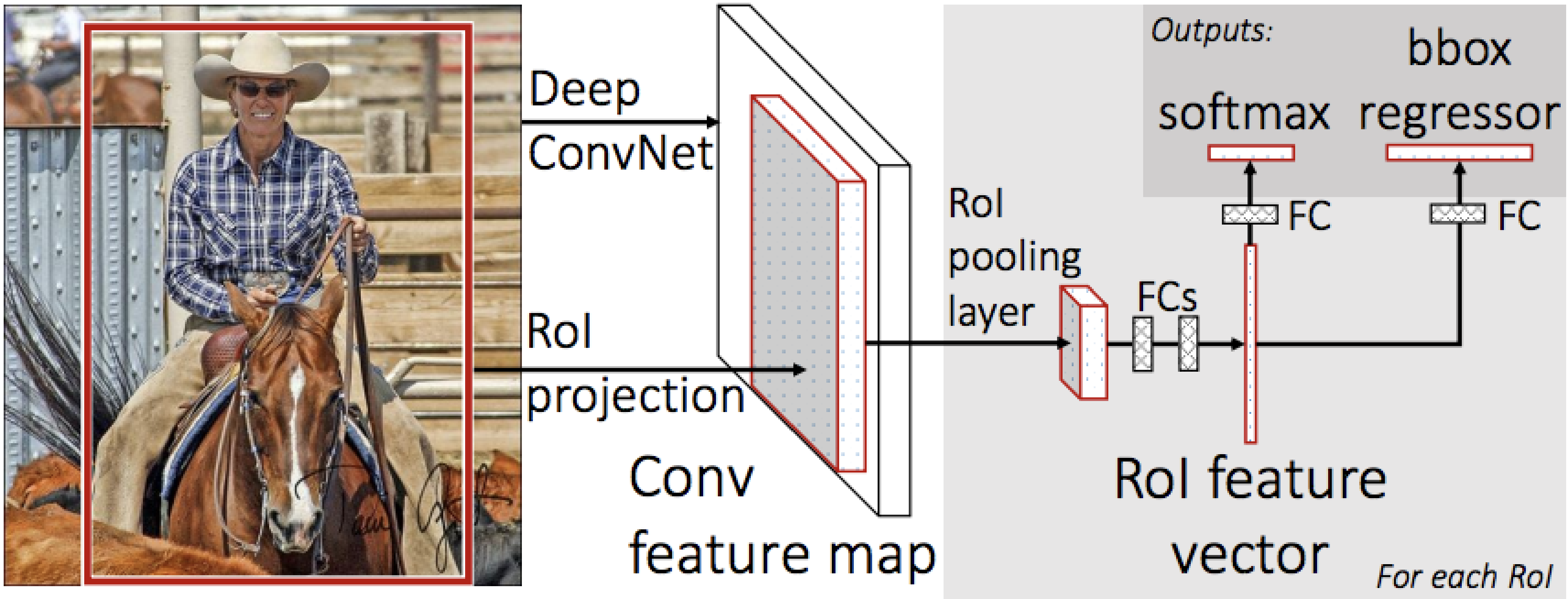
R-CNN



R. Girschick et al. (2014)

- Extract region proposals (RoI) via selective search
- Classify regions with SVM

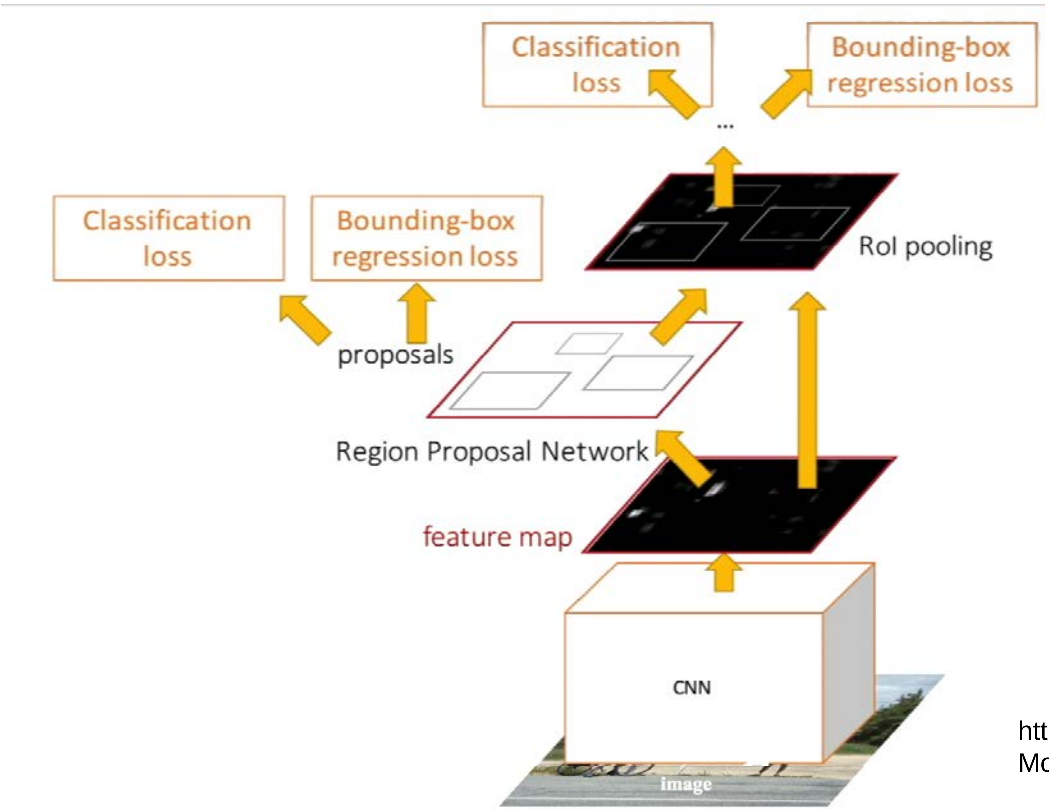
Fast R-CNN



R. Girschick et al. (2015)

- Extract proposals via selective search
- Extract features and classify with CNN

Faster R-CNN

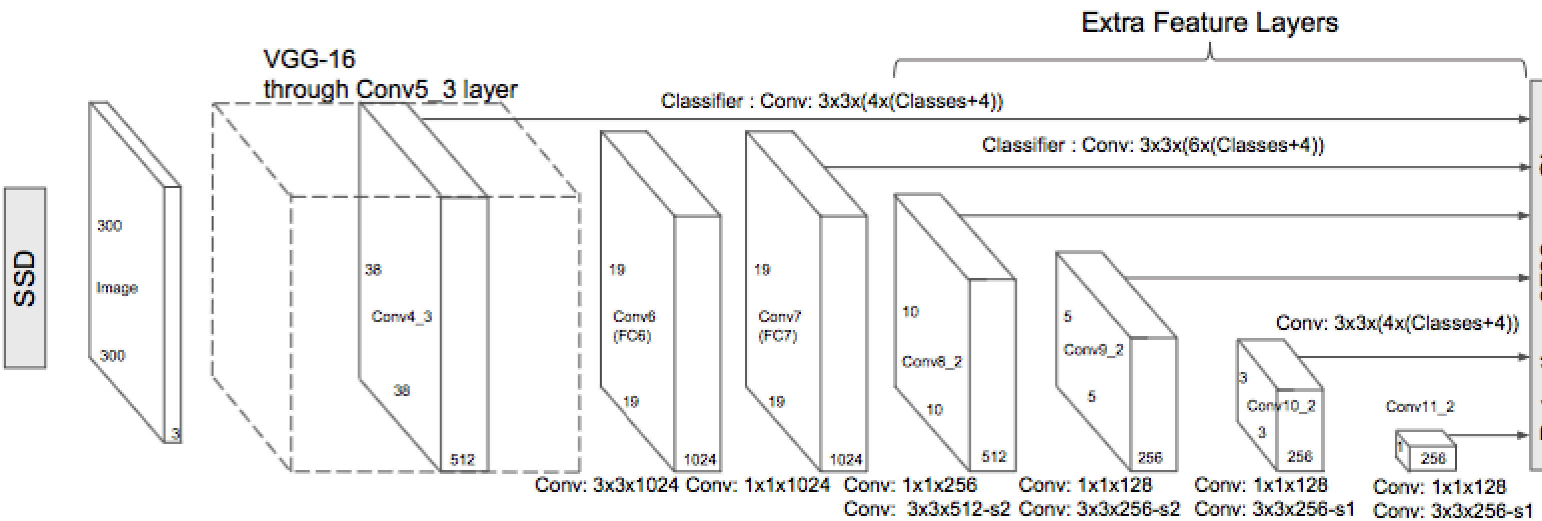


http://cs231n.stanford.edu/
Modified picture from Sh. Ren et al. (2015)

- Region Proposal Network is introduced
- Extract proposals, compute features and classify with CNN

Jointly train with 4 losses

SSD



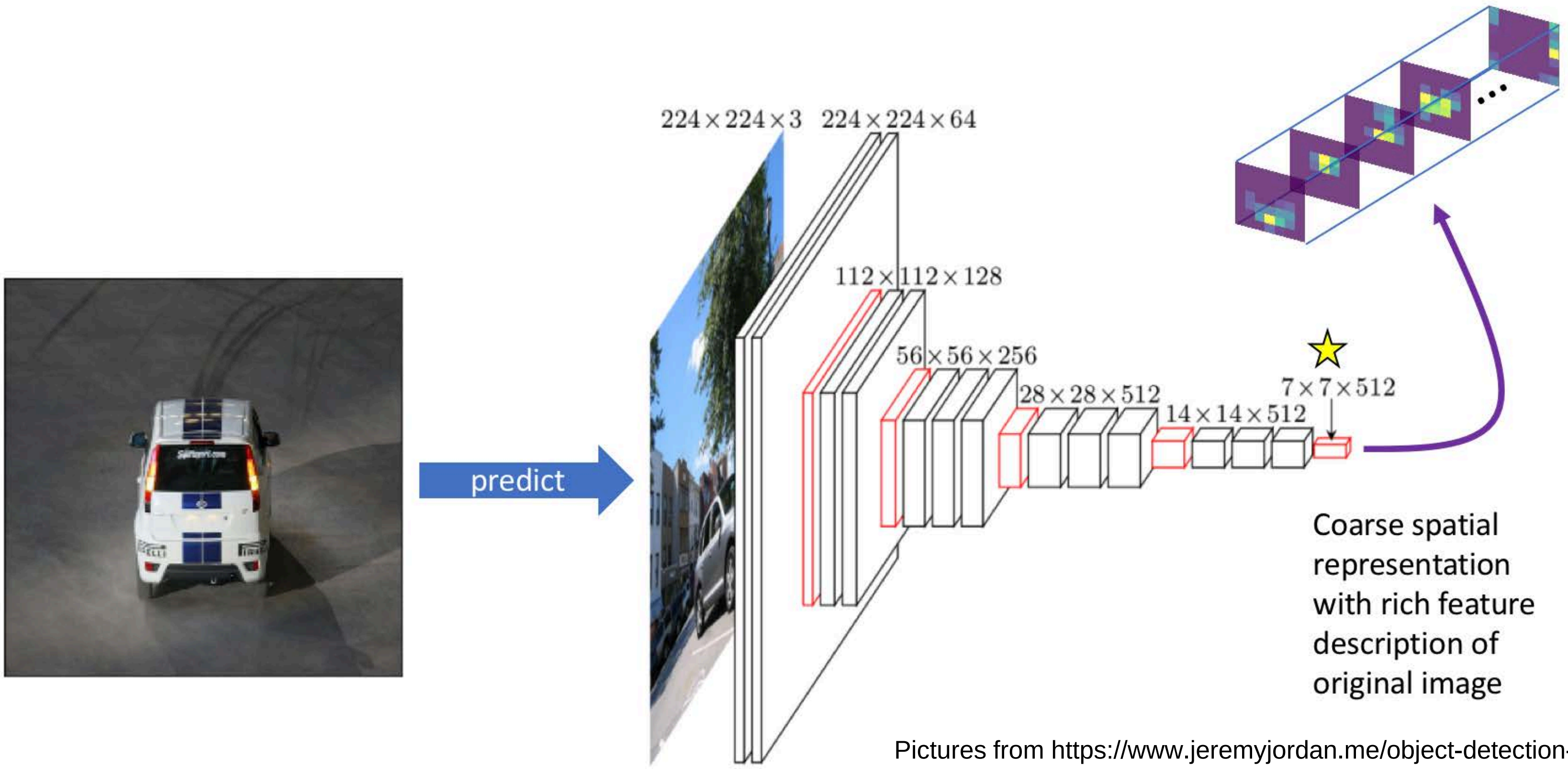
W. Liu et al. (2016)

- Predict class and bounding box at each feature scale

YOLO

YOLOv2 principle

Coarse spatial representation

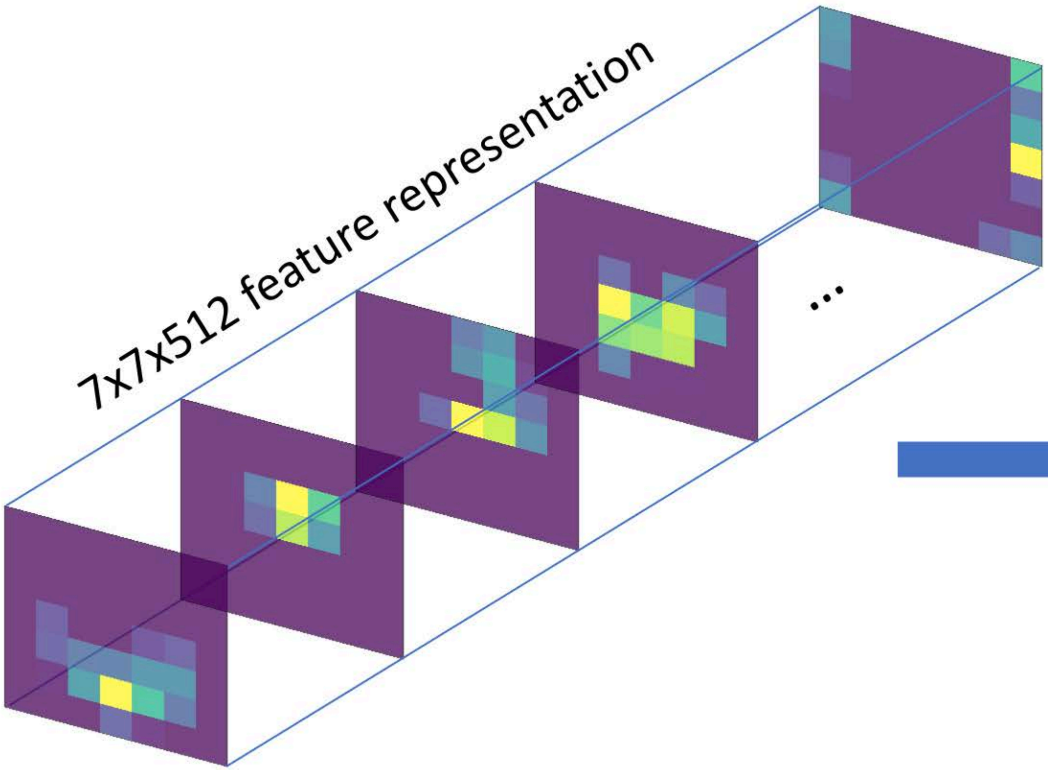


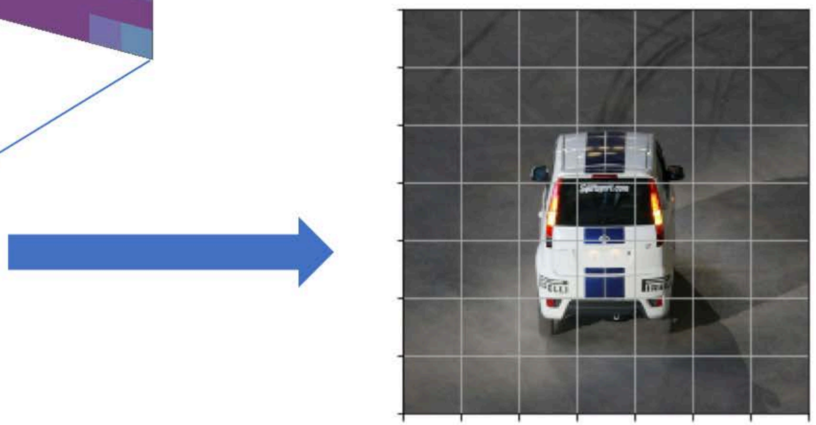
Pictures from https://www.jeremyjordan.me/object-detection-one-stage/

“Backbone” – pretrained CNN for feature extraction

YOLOv2 principle

Coarse spatial representation

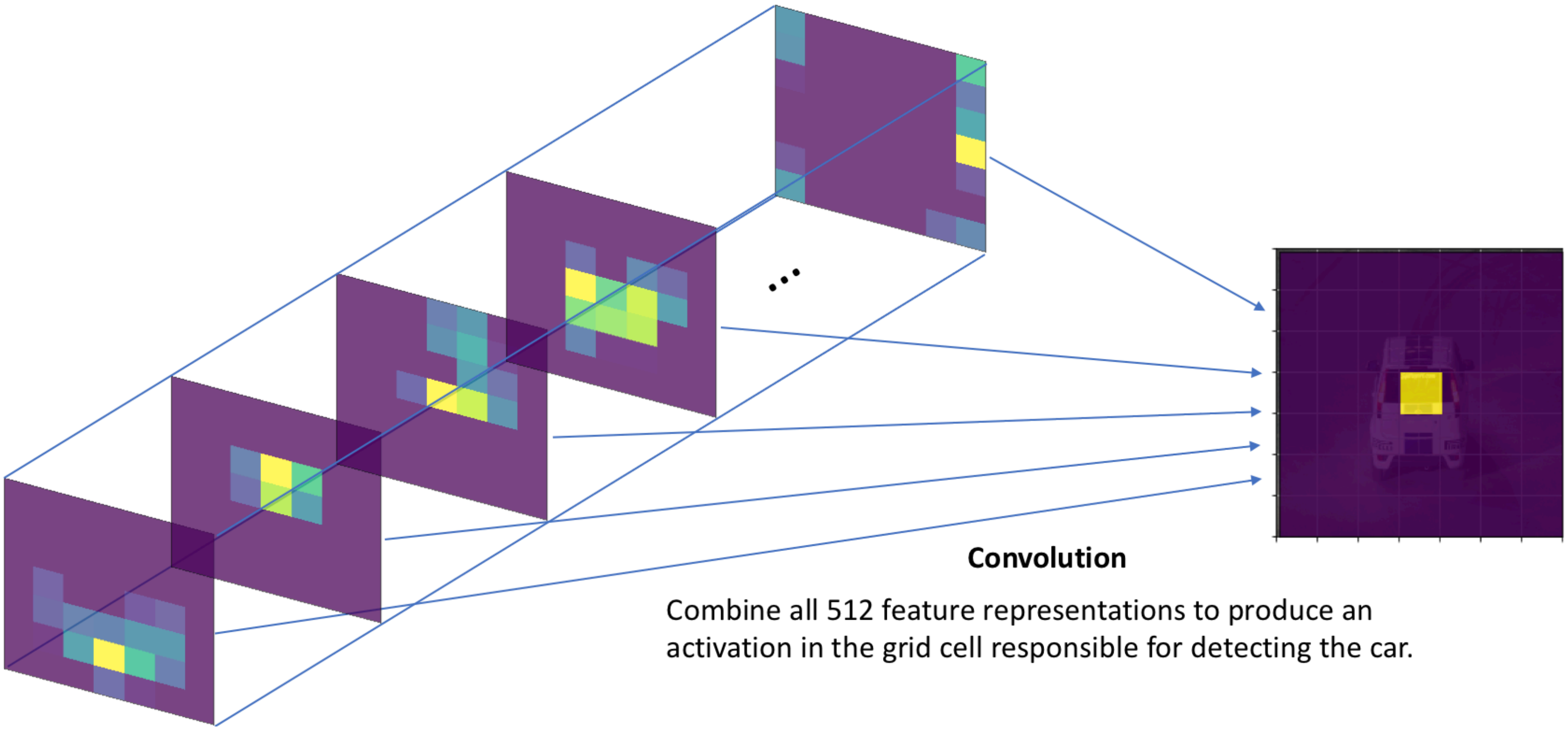




Visualizing the corresponding regions of each "pixel" in the 7x7 feature maps with the original 224x224 image

YOLOv2 principle

Rough object center



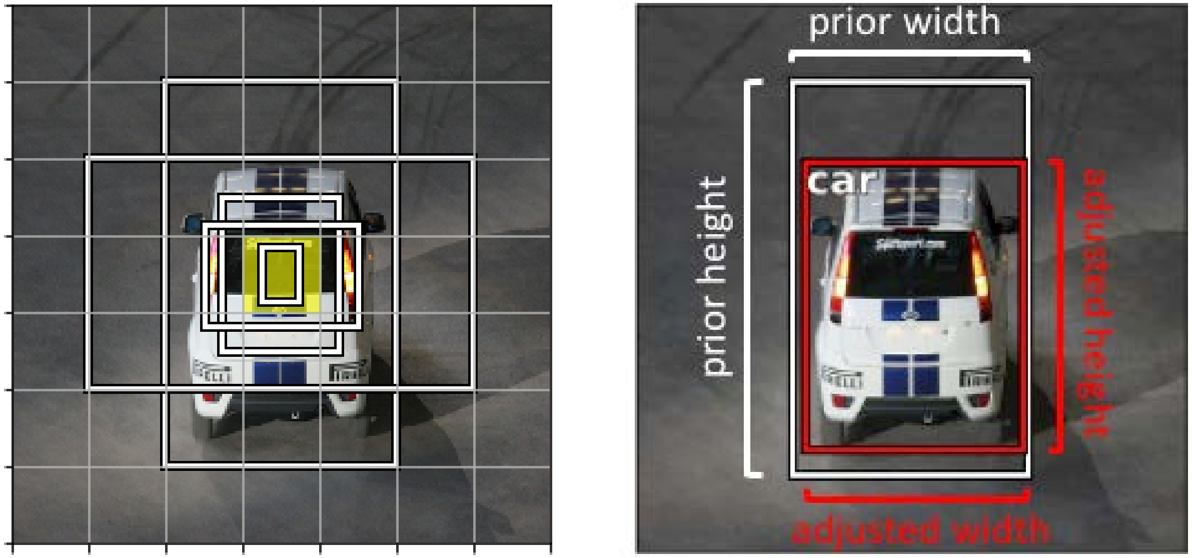
Convolution

Combine all 512 feature representations to produce an activation in the grid cell responsible for detecting the car.

o

YOLOv2 principle

Anchor boxes

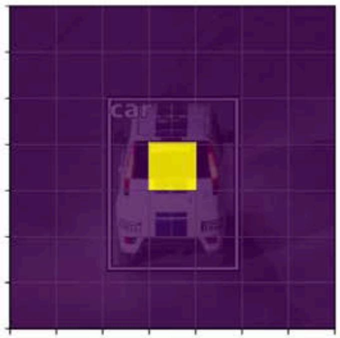


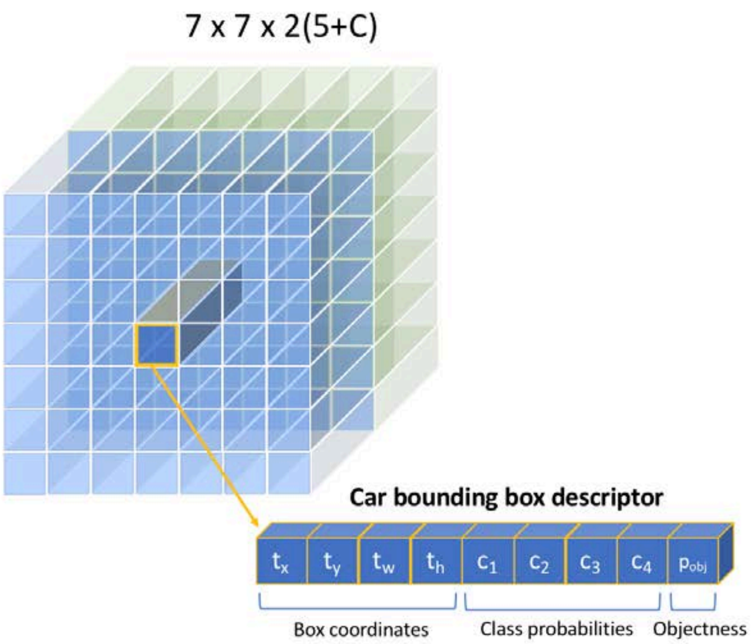
5 bounding box priors (anchor boxes) for each grid cell. Each of 5 bounding boxes specializes in detecting objects of a specific size and aspect ratio

YOLOv2 predicts offsets to each of the anchor boxes instead of predicting arbitrary bounding boxes

YOLOv2 principle

Bounding box descriptor





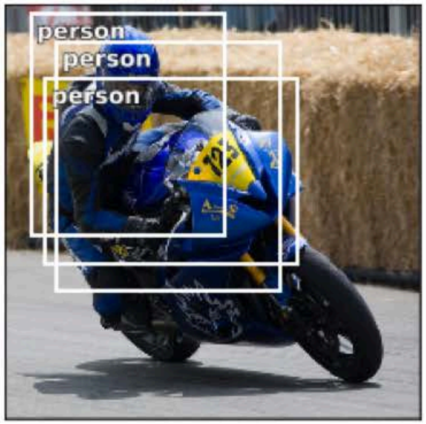
Filter the predictions to only consider bounding boxes which has a p_{obj} above some defined threshold

YOLOv2 principle

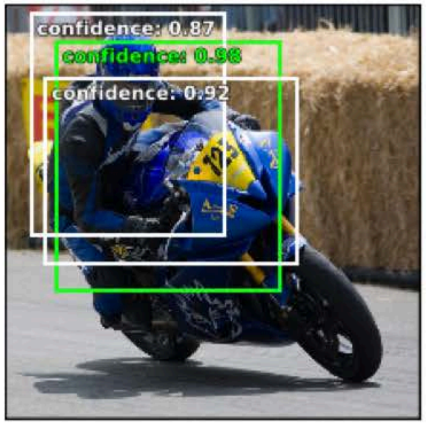
Non-max suppression

Repeat with next highest confidence prediction until no more boxes are being suppressed

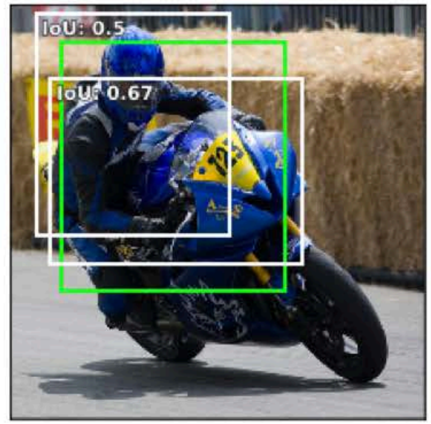
For each class…



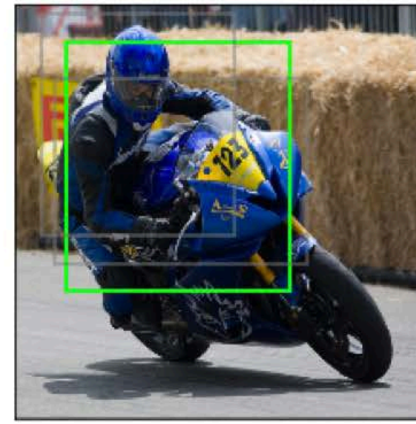
After filtering out low confidence predictions, we may still be left with ***redundant* detections**



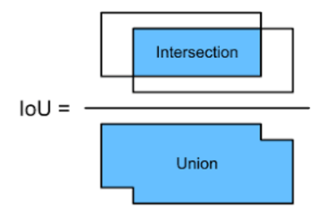
Select the bounding box prediction with the **highest confidence**



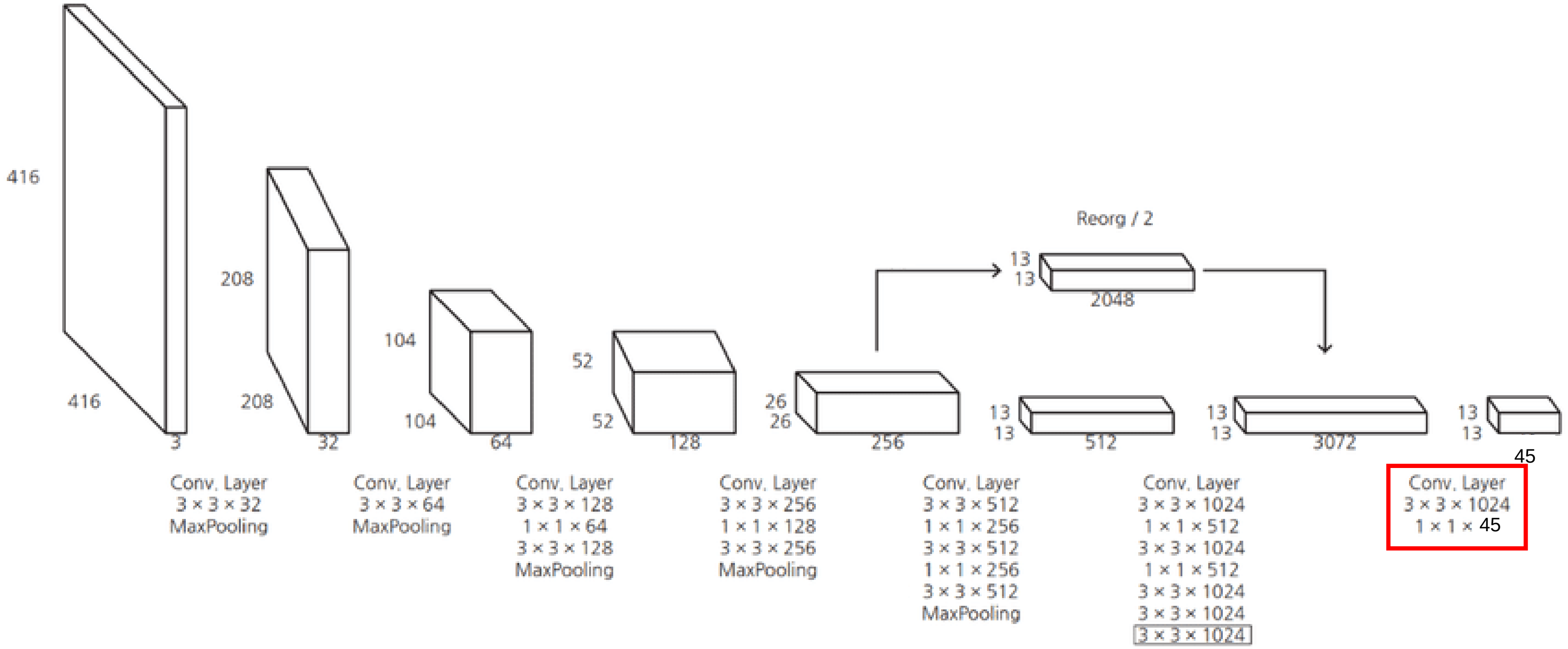
Calculate the IoU between the **selected box** and all remaining predictions



Remove any boxes which have an IoU score above some defined threshold



YOLOv2



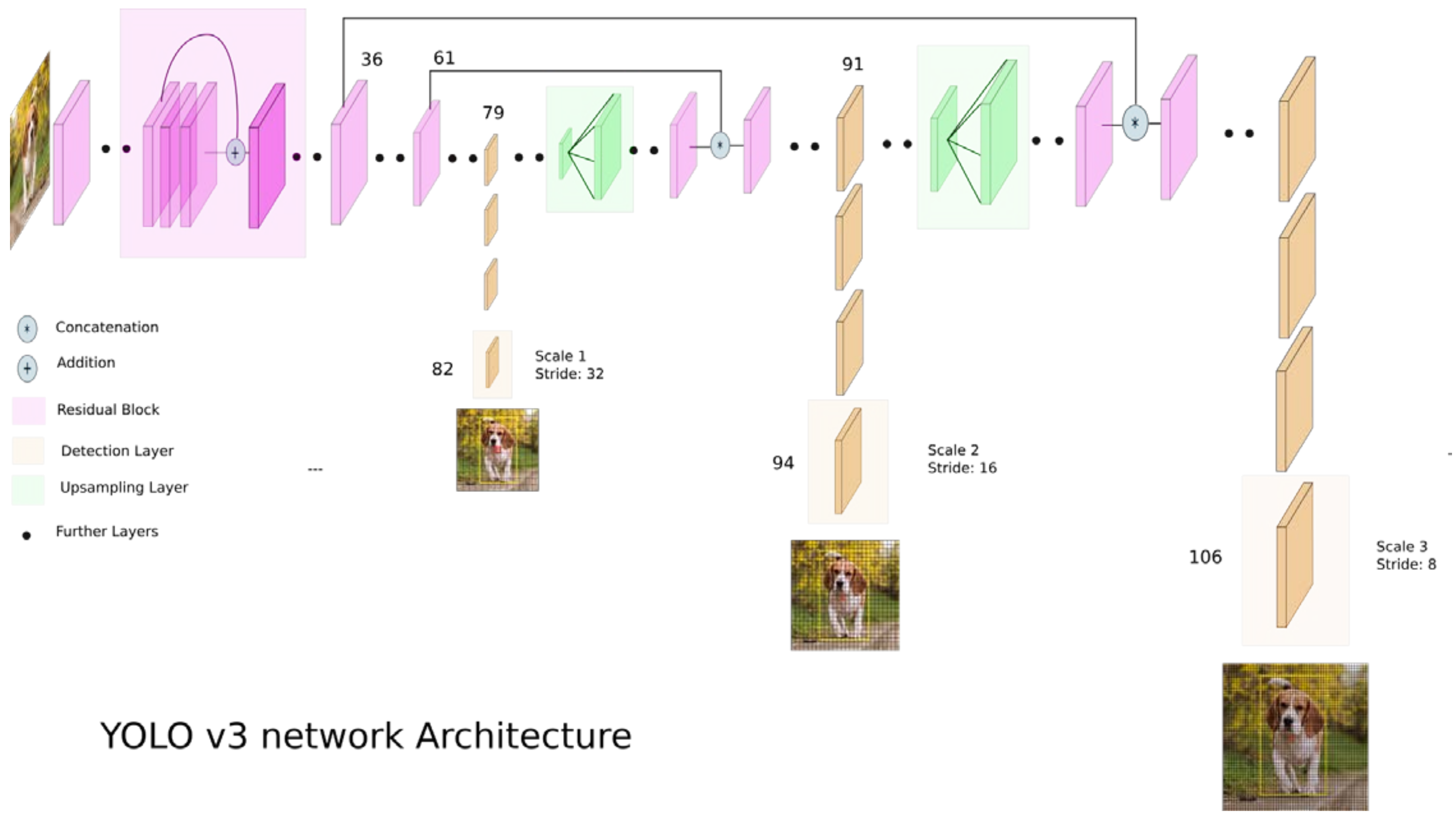
Add 2 convolutional layers to make S×S ×B(5+C) predictions

S=13

B=5

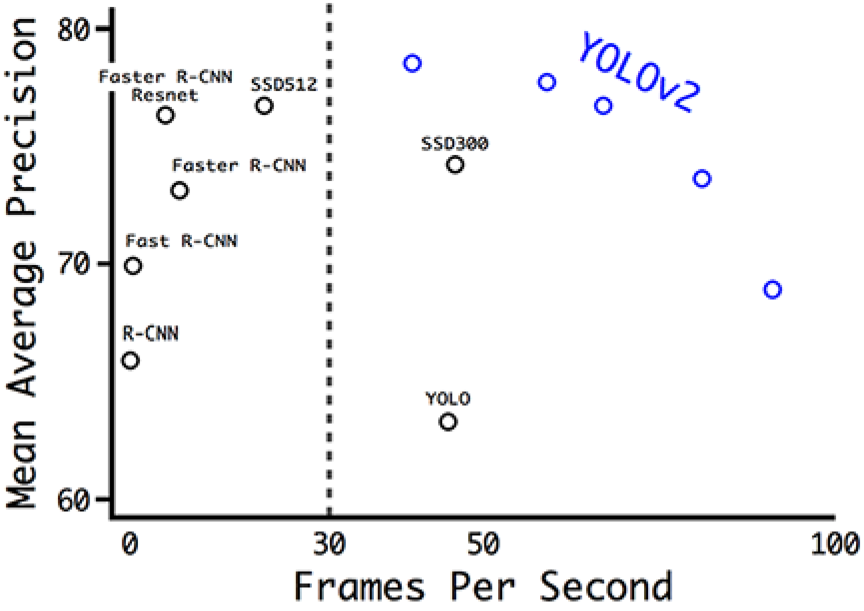
C=4

YOLOv3



- Residual network
- Detection at multiple scales

YOLO compared to other object detectors

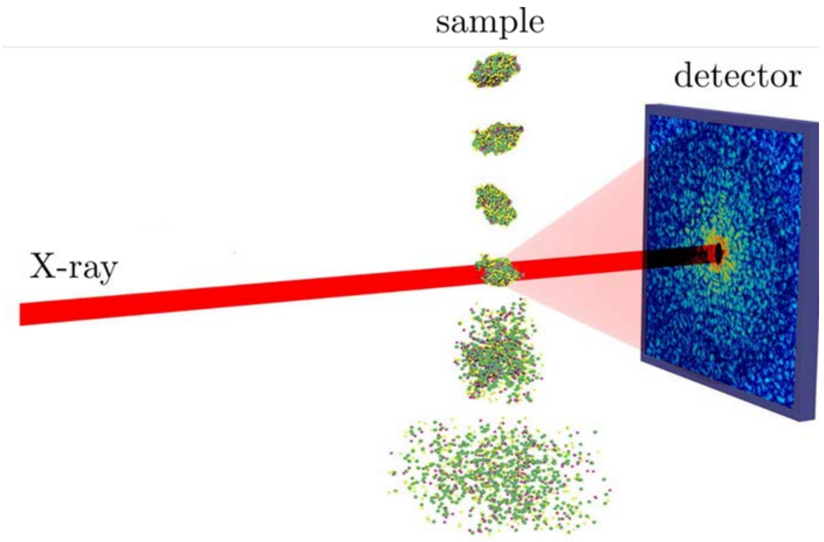


http://pjreddie.com/yolo9000/

Application for classification of diffraction patterns in SPI experiments

Single Particle Imaging (SPI)

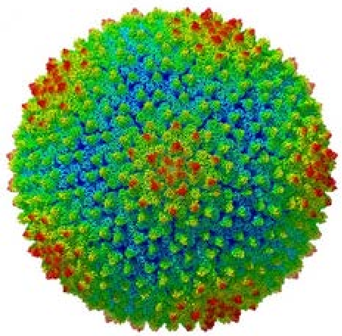
- SPI is a method for native structure determination
- Particles are injected and delivered into X-ray beam
- Forward scattering wave front propagates and is recorded by detector
- Large amount of data is collected



Data pre-processing → filtering → classification → object structure reconstruction

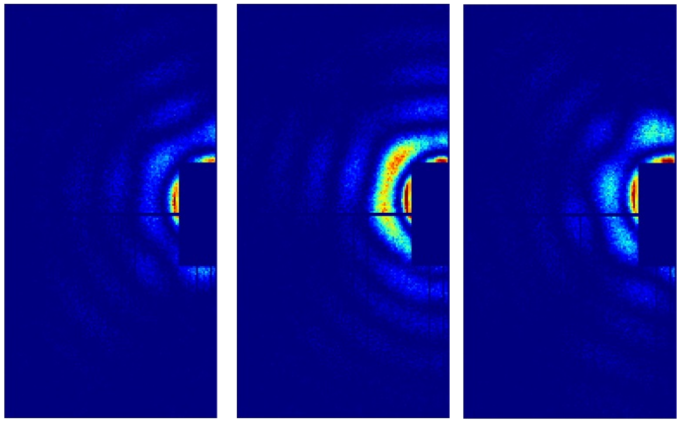
SPI experiment @ LCLS in 2018

- AMO beamline at LCLS
- Sample – bacteriophage PR772, expected size 60-75 nm
- E = 1.7 keV (λ = 7.29 Å)
- Sample-detector distance = 125 mm
- Detector – pnCCD, a half of it was operational

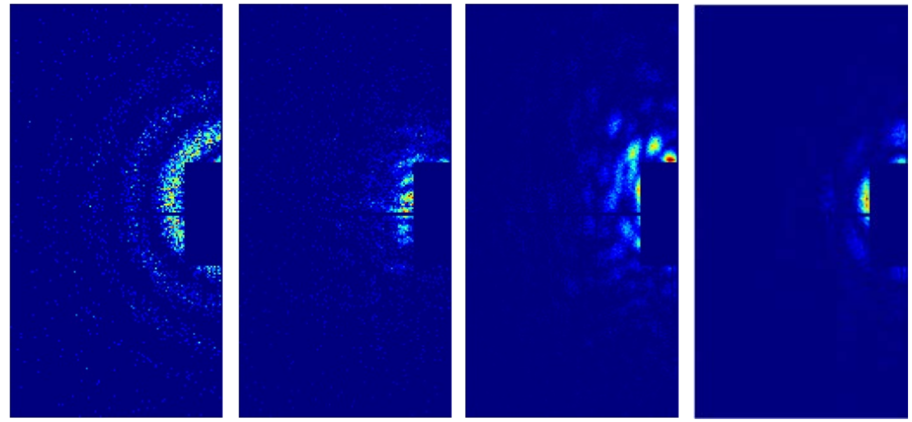


SPI experiment @ LCLS in 2018

Positive examples:



Negative examples:



Subtracted background, photon counts

Training and validation sets

Training set

Positive set

- 165 patterns in the size range 60 -75 nm

Negative set

- 373 carefully selected patterns across all runs

Validation set

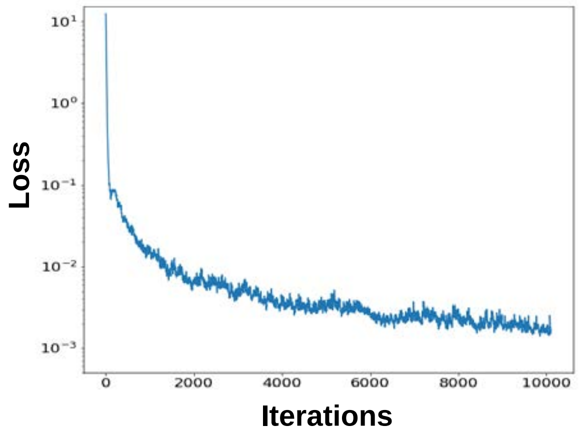
Positive set

- 53 patterns in the size range 60 -75 nm

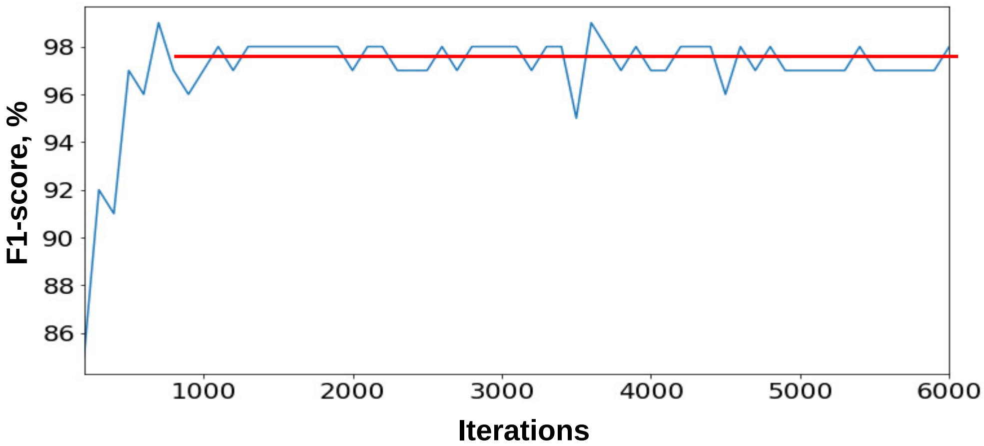
Negative set

- 200 carefully selected patterns across all runs

Training and model choice



Typical training loss



Apply model with the saturated F_1 score

Metrics

$$F_\beta = (1+\beta^2)\cdot\frac{\text{precision}\cdot\text{recall}}{(\beta^2\cdot\text{precision})+\text{recall}}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

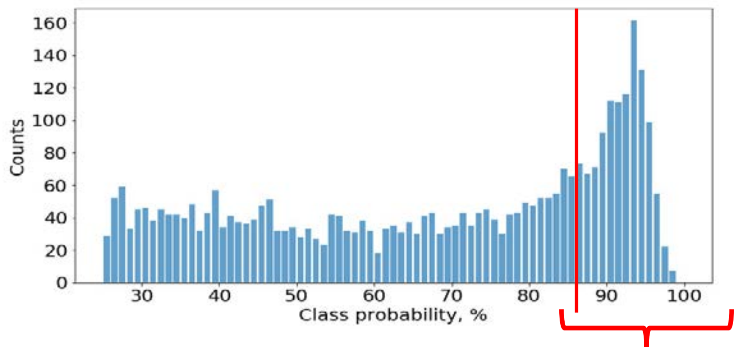
$$F_1 = 2\cdot\frac{\text{precision}\cdot\text{recall}}{\text{precision}+\text{recall}}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

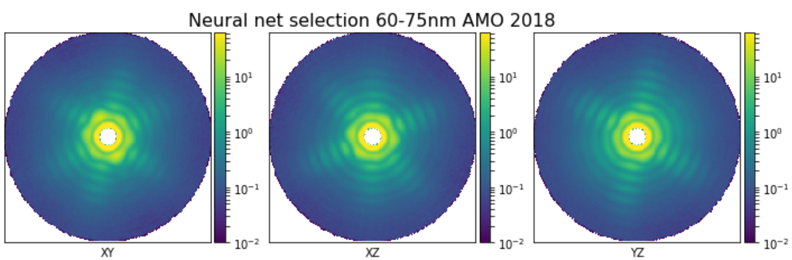
Results

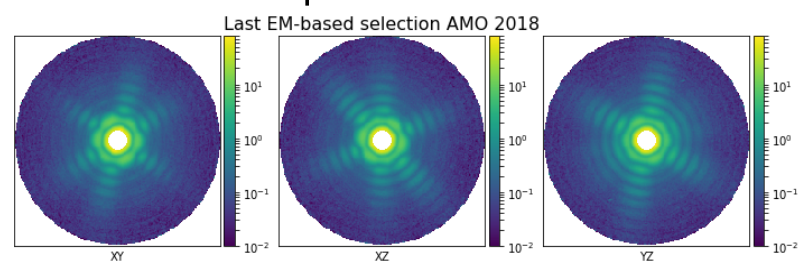
Model run on ~18k patterns filtered by particle size

86%



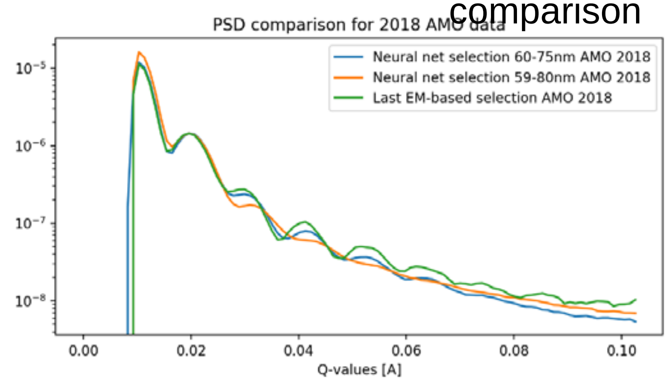
1036 patterns selected





EMC reconstruction with 1036 patterns

EMC reconstruction with EM-based selection for comparison



Summary & outlook

- CNN-based selection of single hits is promising – it was possible to make reasonable selection
- Further steps:
  - Supplement training and validation data with simulated examples
  - Use residual network
  - Use intensity values instead of 3D color scheme as an input

Acknowledgements

Coherent X-ray Scattering and Imaging Group at DESY

Prof. Dr. I. A. Vartanyants
N. Mukharamova
D. Assalauova
D. Lapkin
R. Khubbutdinov

Dr. J. Carnis
Dr. Y-Y. Kim
Dr. A. Ignatenko
Dr. L. Gelisio
Dr. M. Rose

Prof. Dr. E. Weckert (DESY)

S. A. Bobkov (National research Center “Kurchatov Institute”)

SPI consortium @ LCLS

A. Aquila (LCLS)

Helmholtz Associations Initiative and Networking Fund and the Russian Science Foundation grant HRSF-0002/18-41-06001

RSF | Russian Science Foundation

Thank you for your attention