Object detection with Convolutional Neural Nets

Falgayrac Jérôme, Hardouin Baptiste, Virel Guillaume, Voreiter Damien

M2 - University of Bretagne Sud, Vannes

2019

Table of contents

Table of contents

Convolutional Neural Networks *(CNN)*

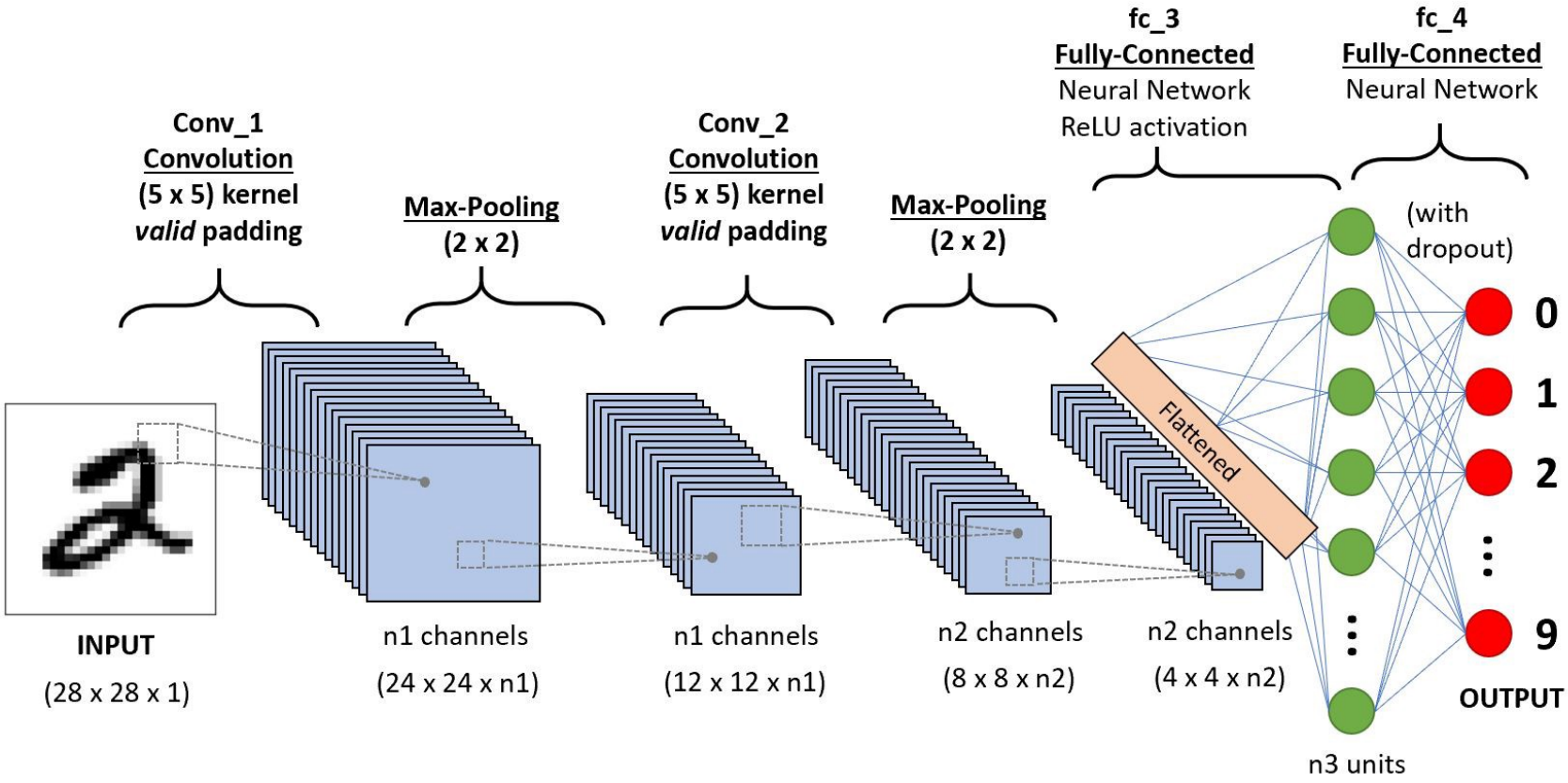


Figure: Convolutional Network

Support Vector Machine *(SVM)*

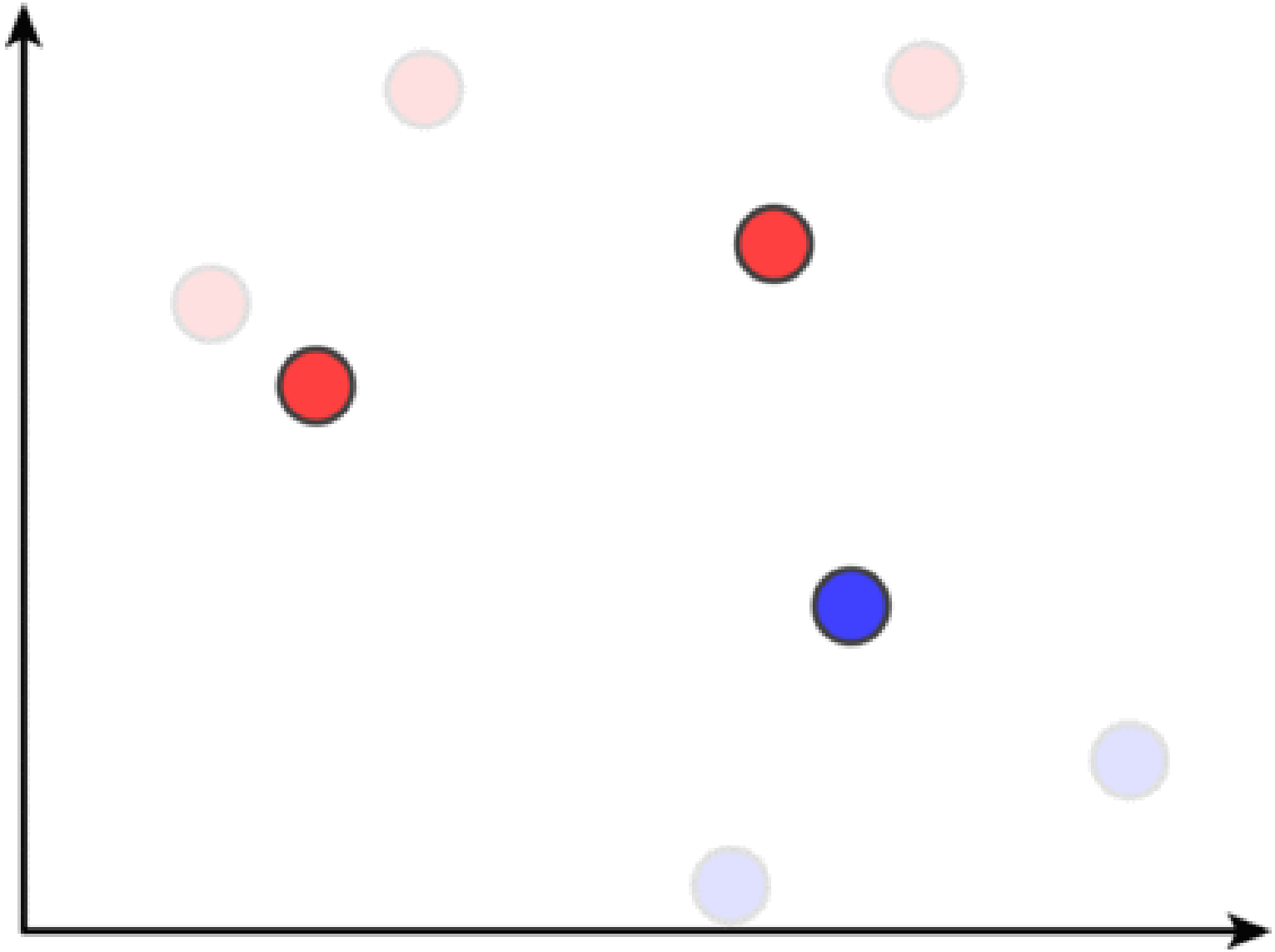
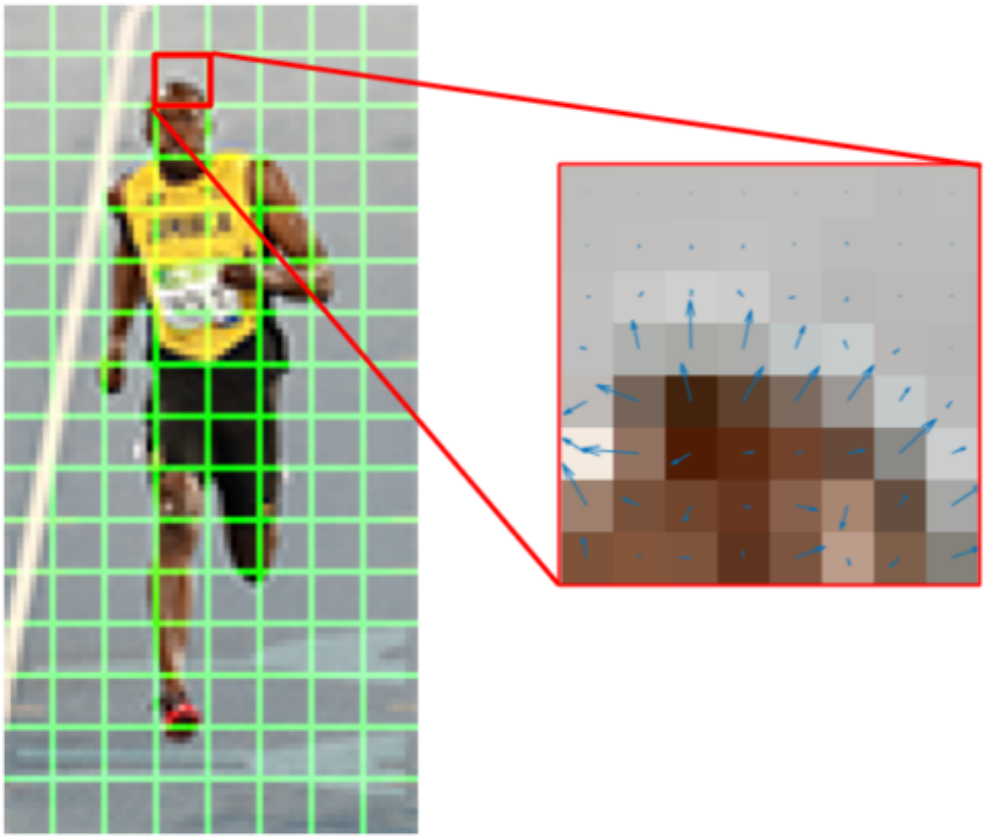


Figure: Working

Table of contents

Histogram Oriented Gradient *(HOG)*



2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	21	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

Gradient Magnitude

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

Gradient Direction

Figure: 8x8 HOG

Histogram Oriented Gradient *(HOG)*

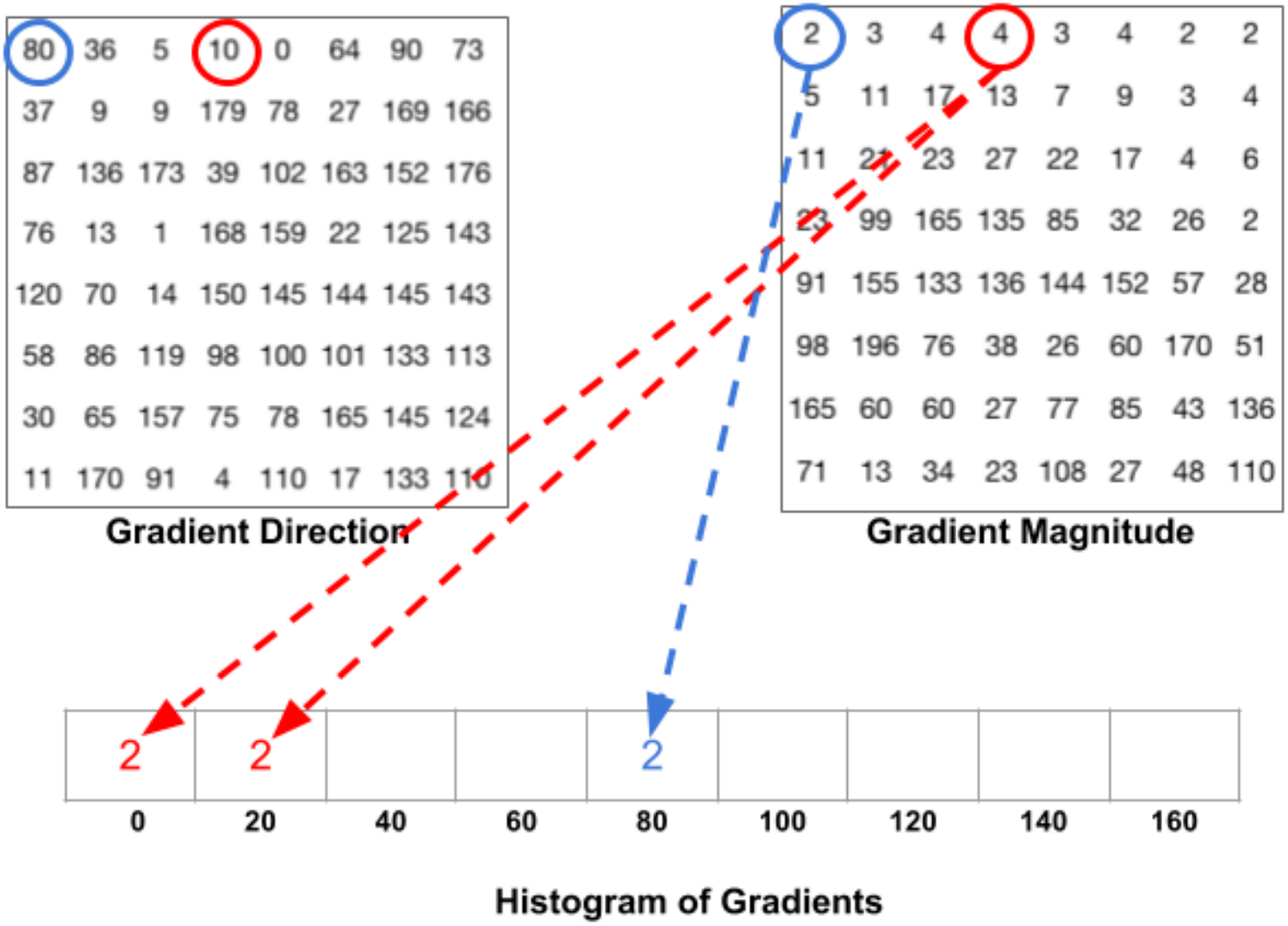


Figure: Fill principle of the histogram

Histogram Oriented Gradient *(HOG)*

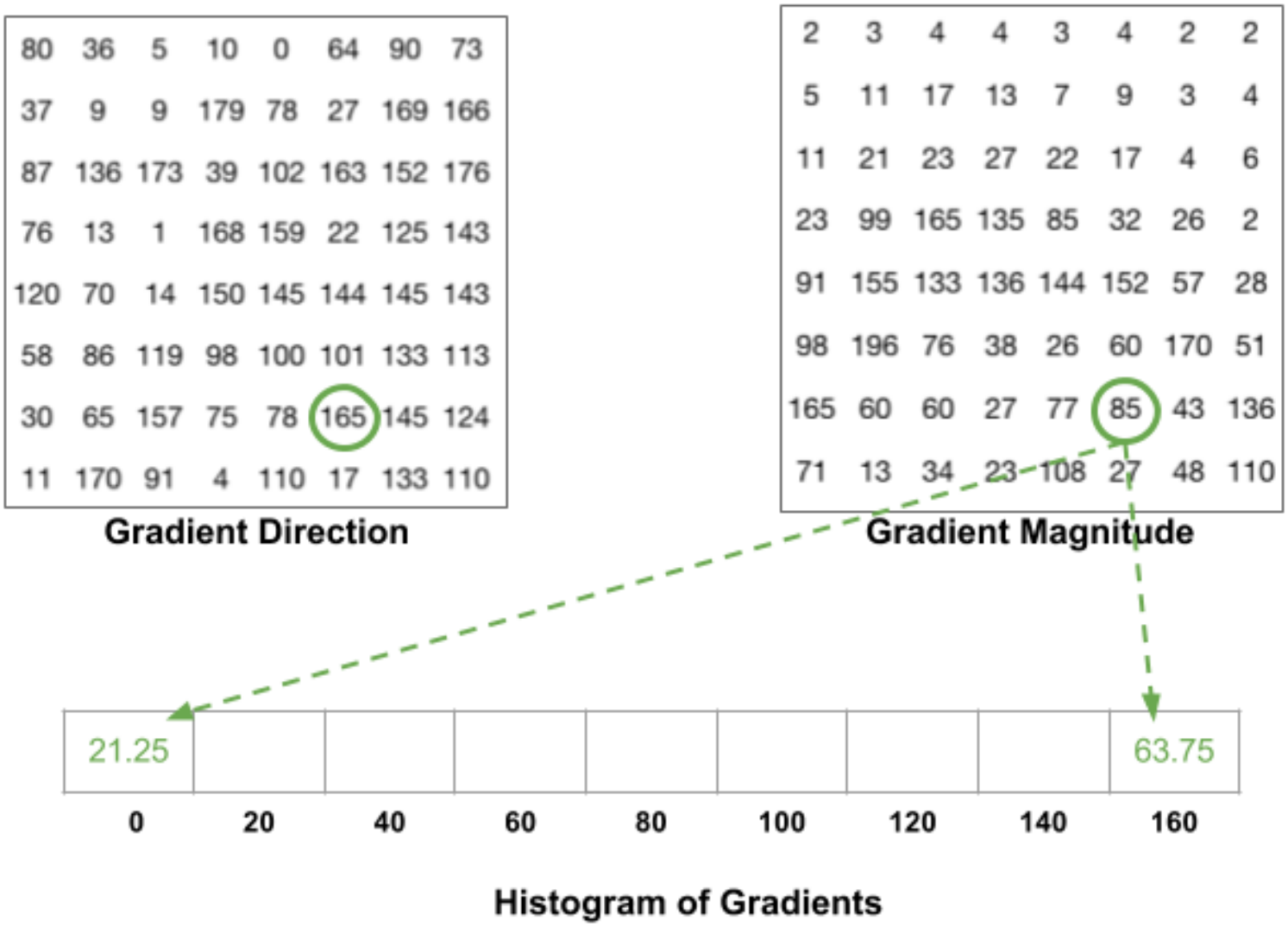


Figure: Fill principle of the histogram

HOG Normalization and result

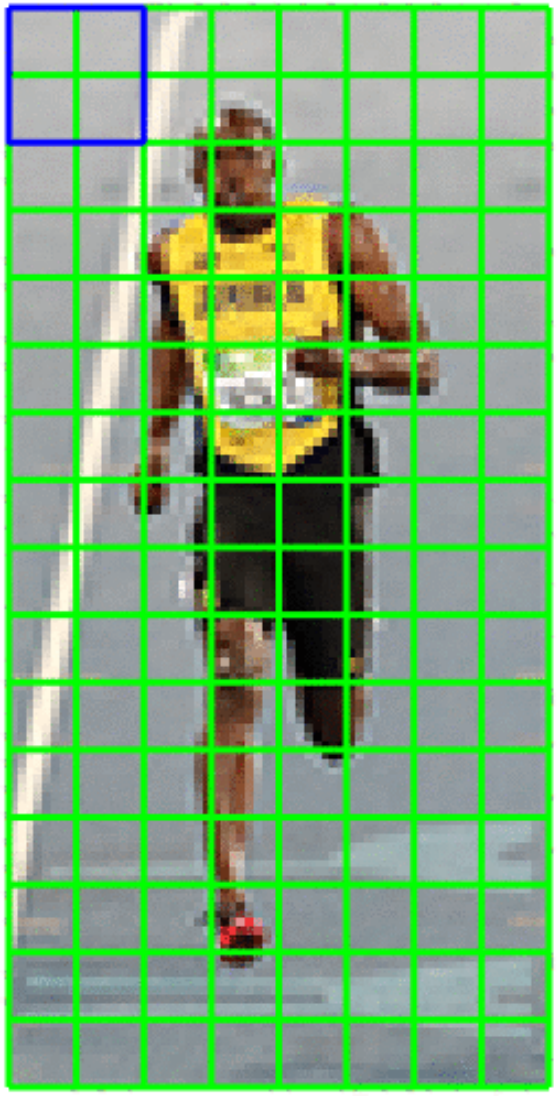


Figure: 16x16 block for normalization

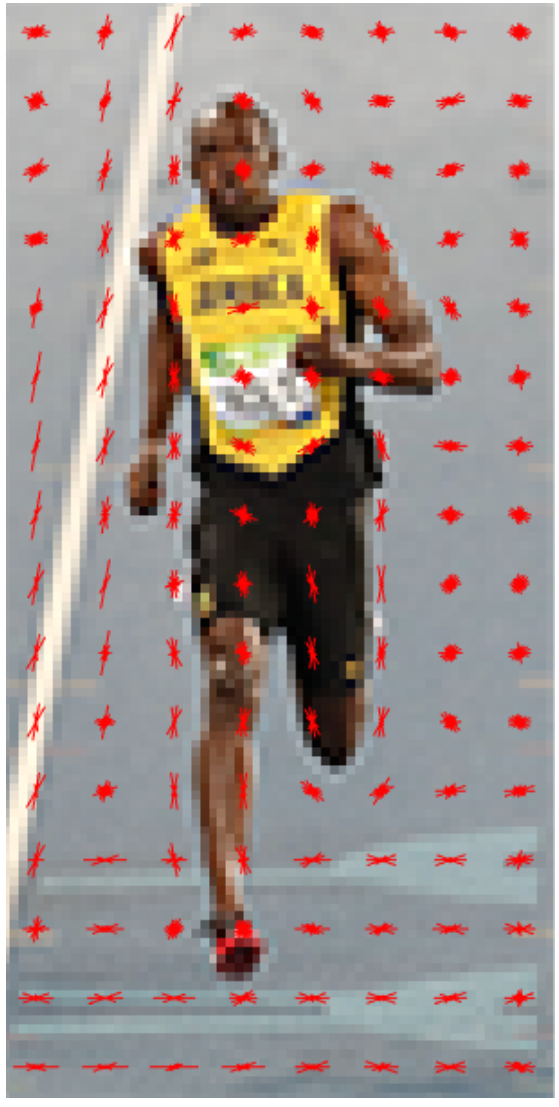


Figure: HOG Visualization

Region-Based Convolutional Neural Networks *(R-CNN)*

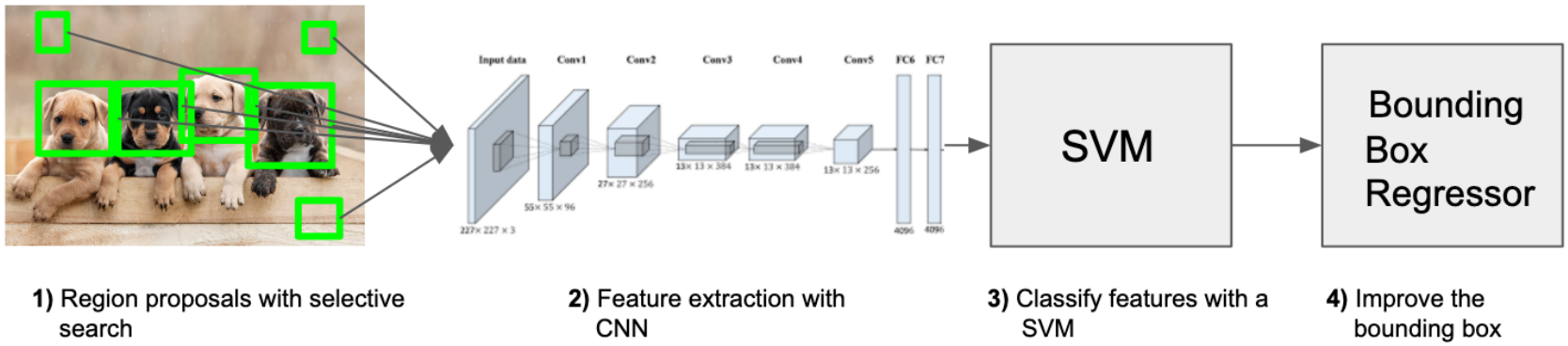


Figure: Model R-CNN

Selective Search

Step 1:

Segmentation method by Felzenszwalb et Huttenlocher

Figure: Input image

Figure: Output image

Step 2:

Figure: Oversegmented images

Similarity

- Color similarity
- Shape compatibility
- Size similarity
- Texture similarity
- Final similarity

Selective Search

Step 3:

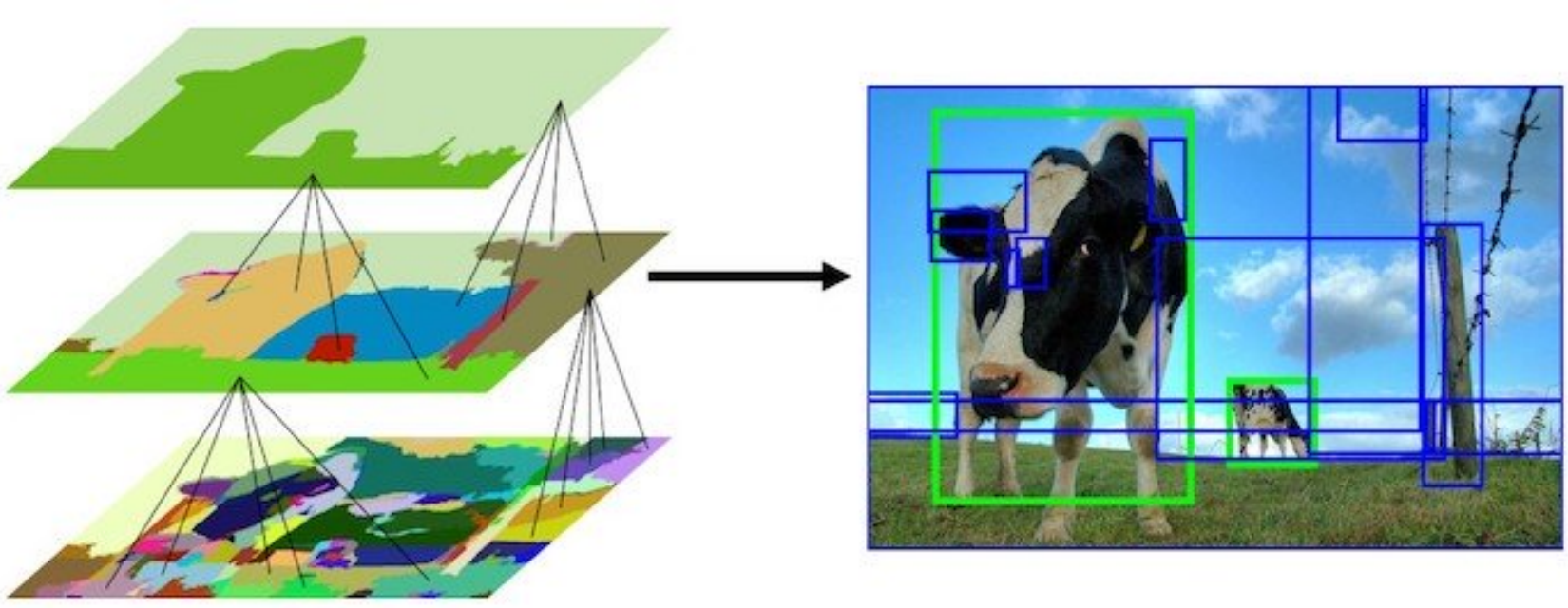


Figure: Grouping similarity regions

Bounding Box Regression

Before non-max suppression

After non-max suppression

Figure: Multiple bounding boxes detect the car in the image

Advanced Methods

Fast R-CNN

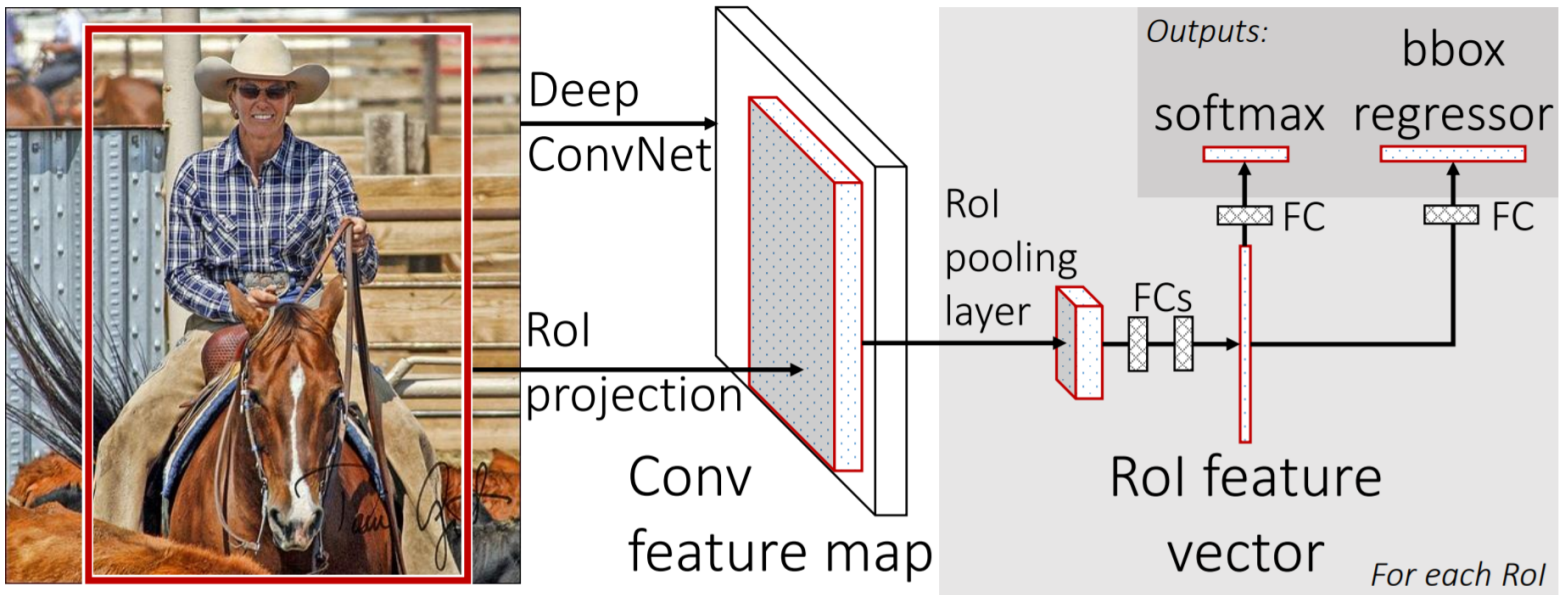


Figure: Fast-RCNN Principle

Faster R-CNN

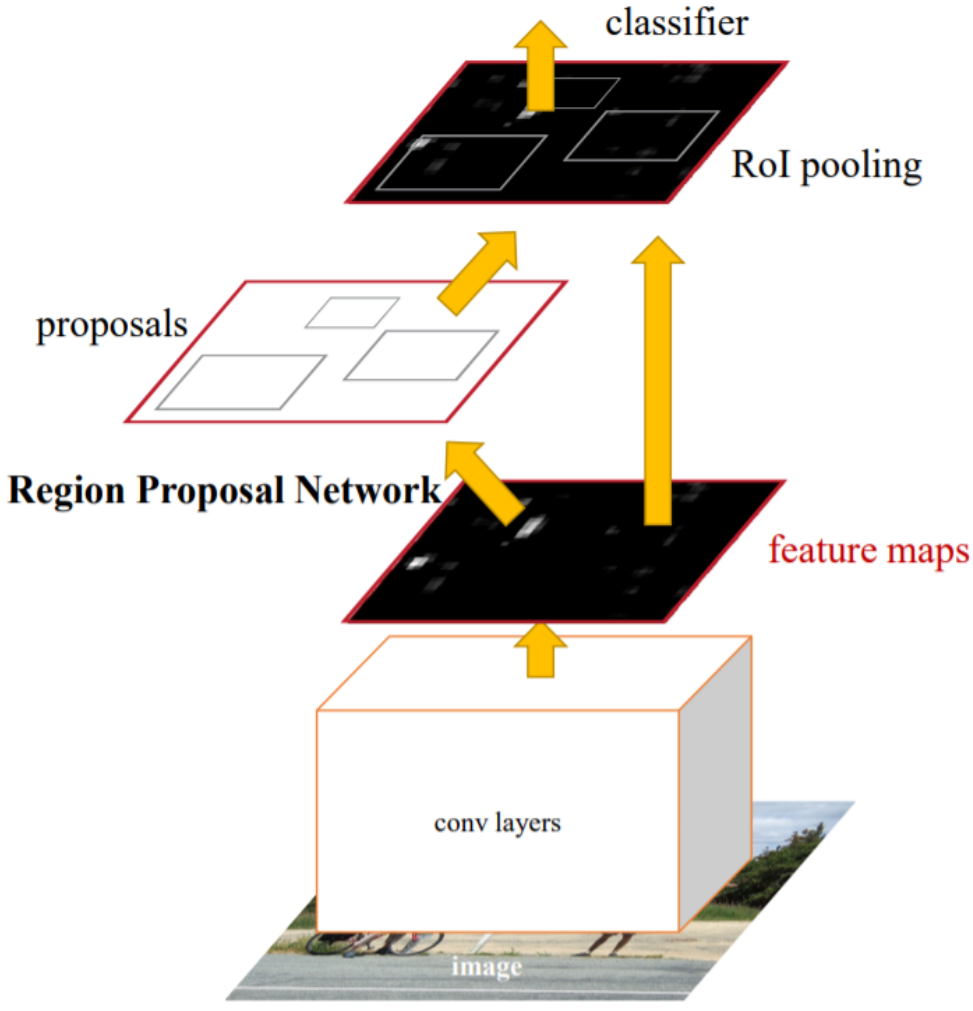


Figure: Faster-RCNN Principle

Region Proposal Network *(RPN)*

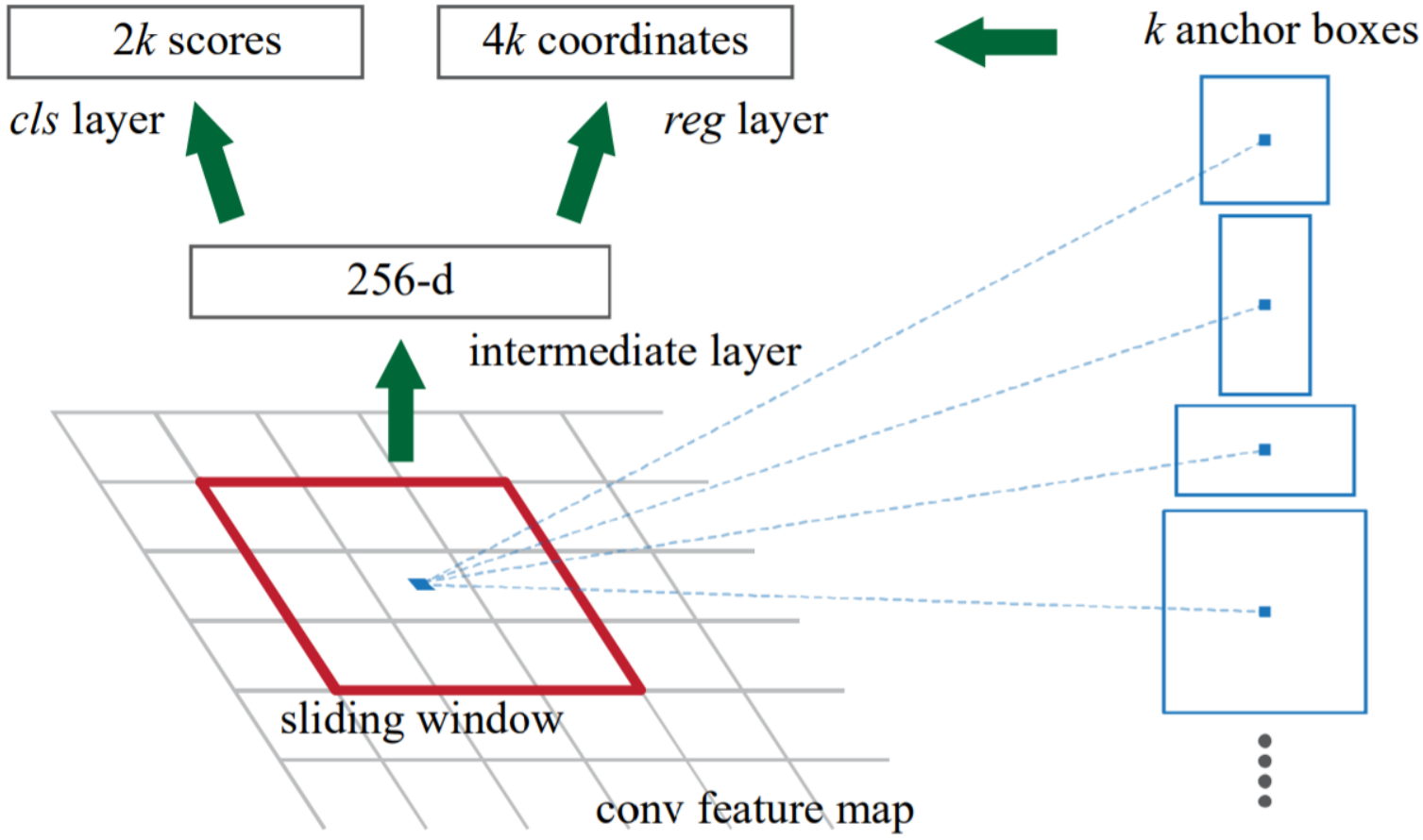


Figure: RPN Principle

Faster-RCNN Results

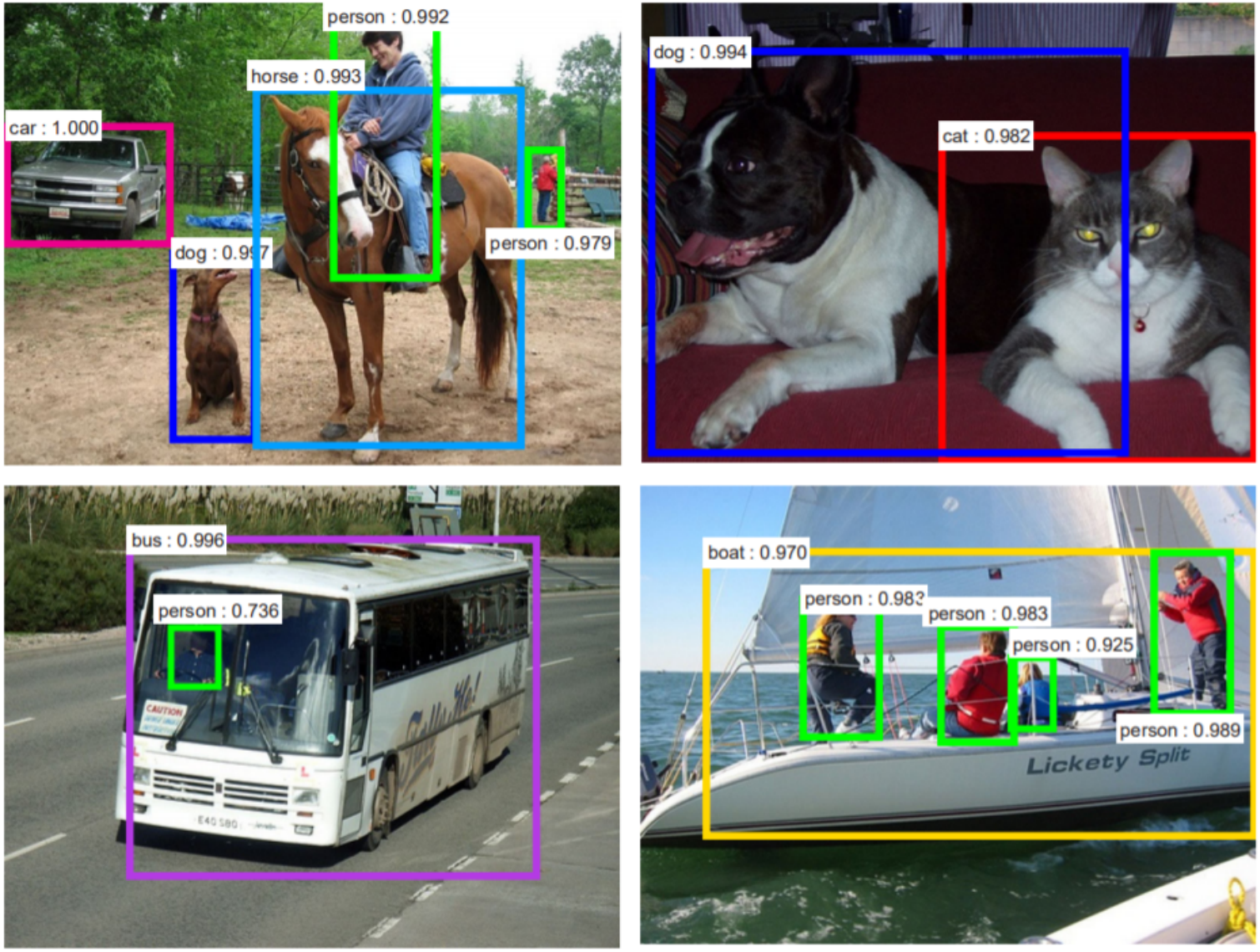


Figure: RPN Principle

YOLO

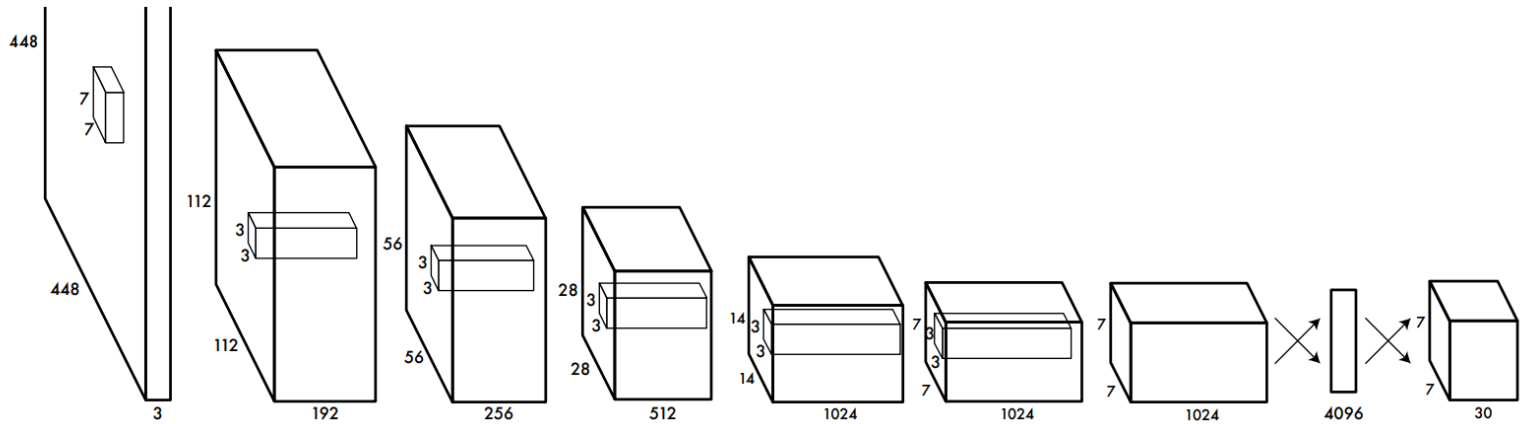


Figure: Yolo Network
(24 + 2 layers)

Pascal VOC 2012 dataset

Figure: Base image

YOLO

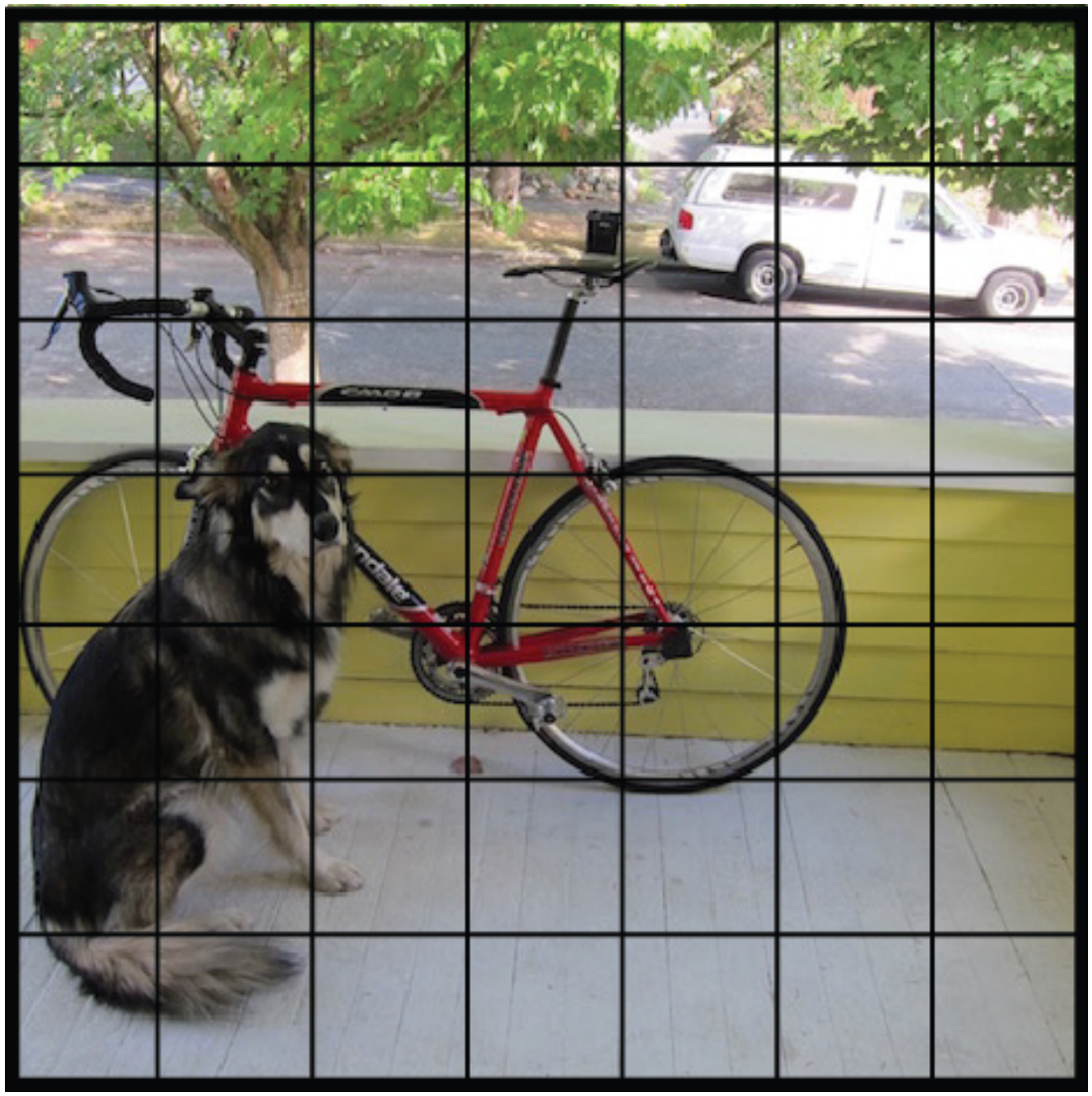


Figure: Image divided into grid (7 x 7)

YOLO

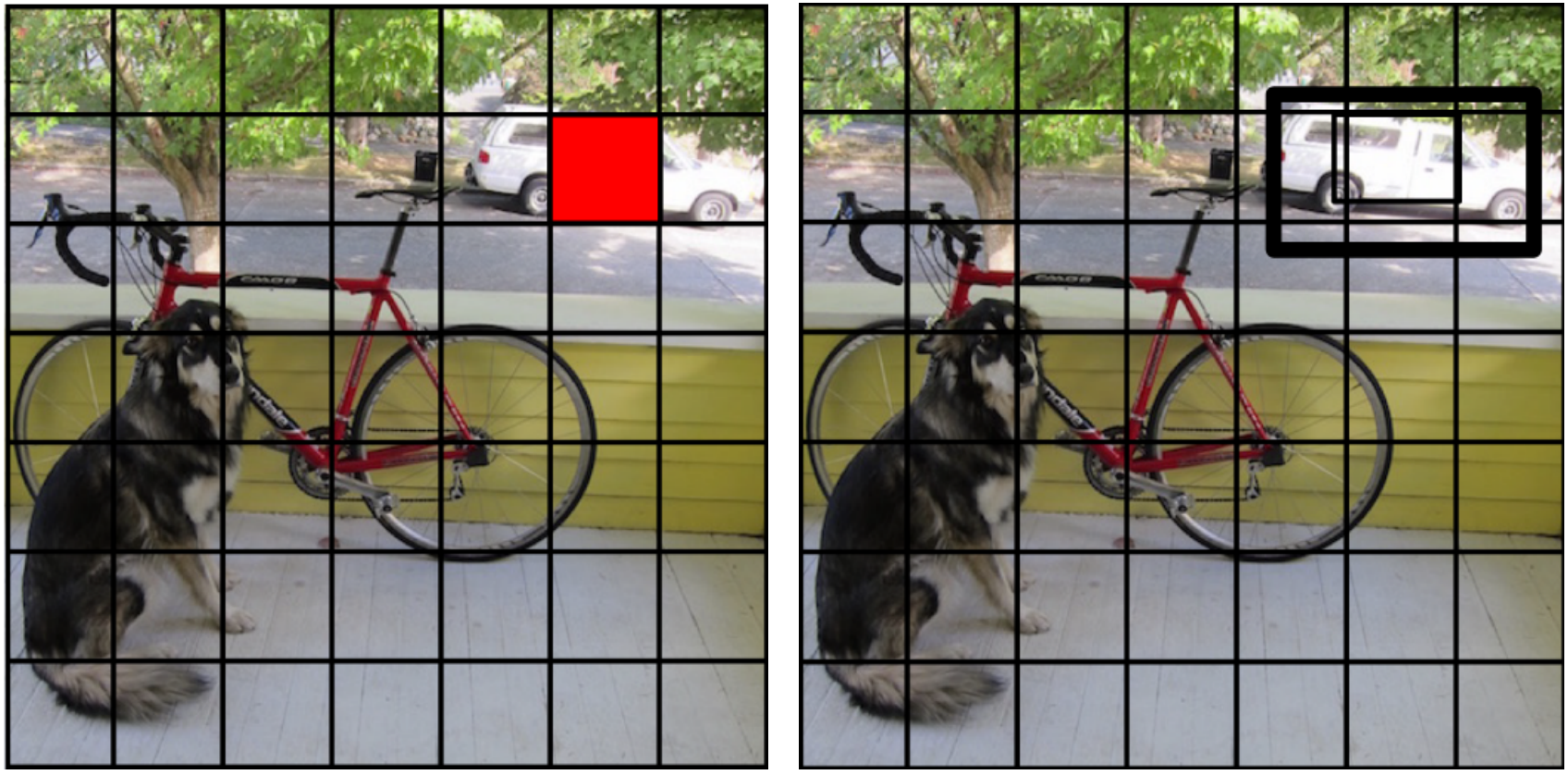


Figure: Exemple on one cell

YOLO

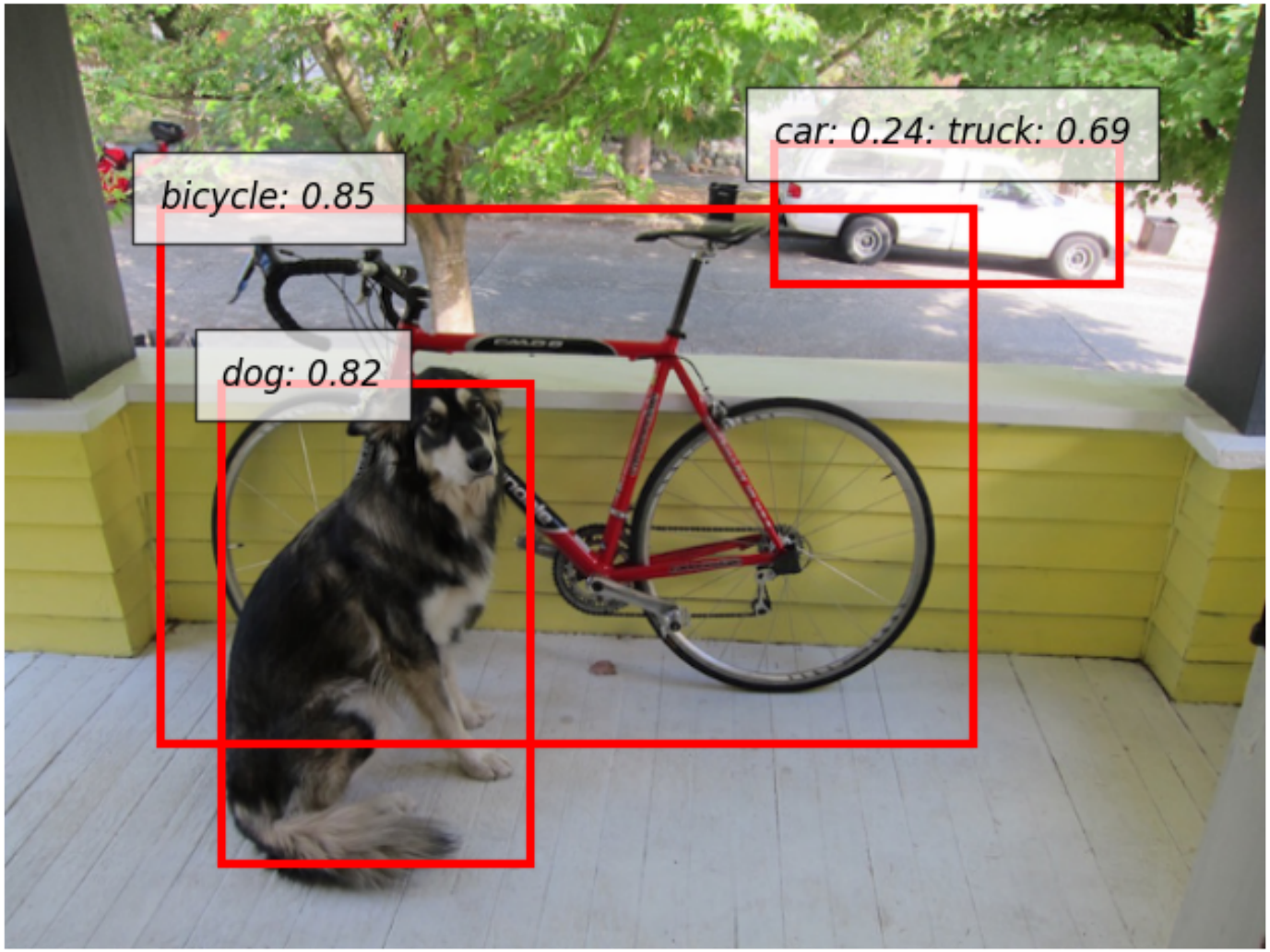


Figure: Result of exemple

Mask R-CNN

Figure: Base image

Mask R-CNN

Figure: Segmentation masks

Mask R-CNN

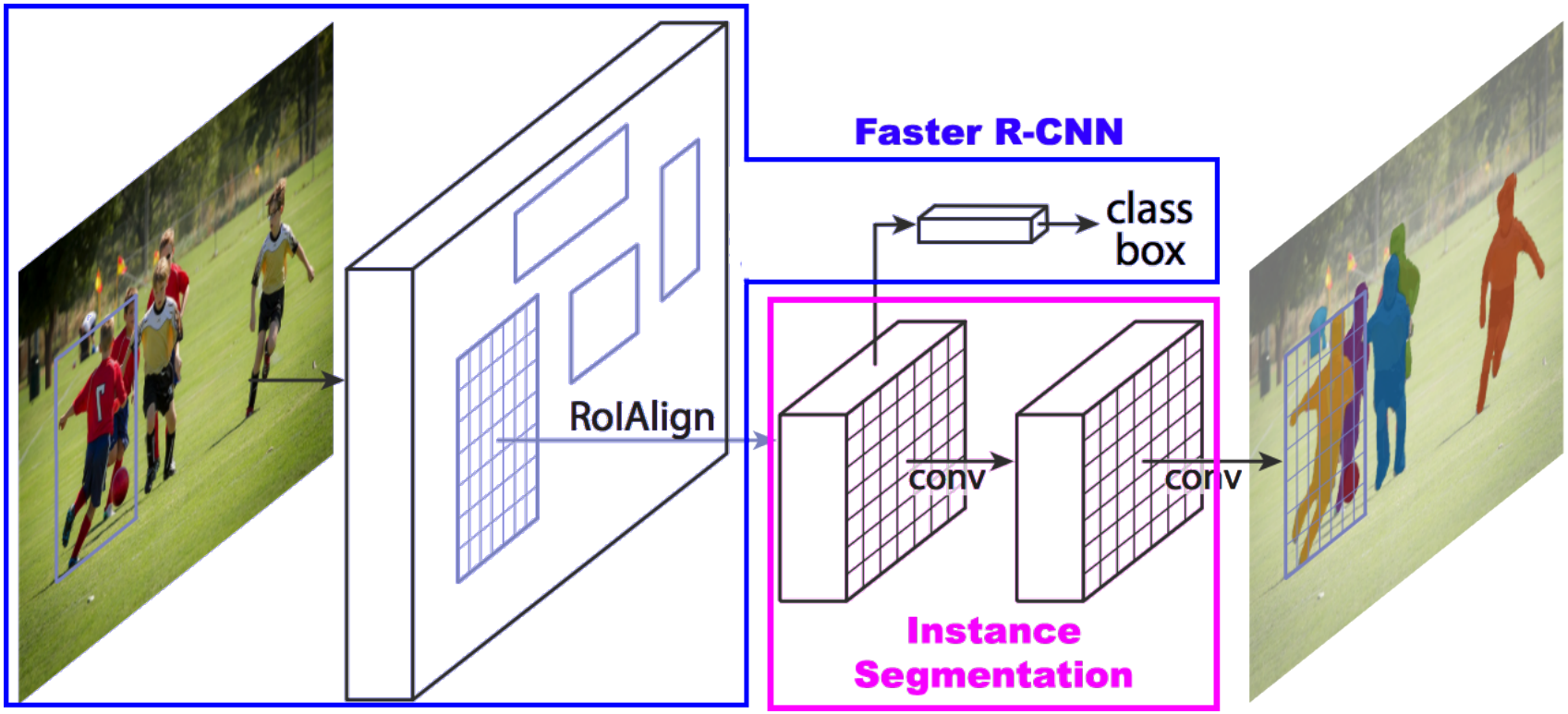


Figure: Mask Model

COCO dataset

Figure: Example

(COCO is only 330 000 images with 1.5 million object instances for 80 classes. . .)

Table of contents

Results

Figure: Accuracy / Speed graph

Results

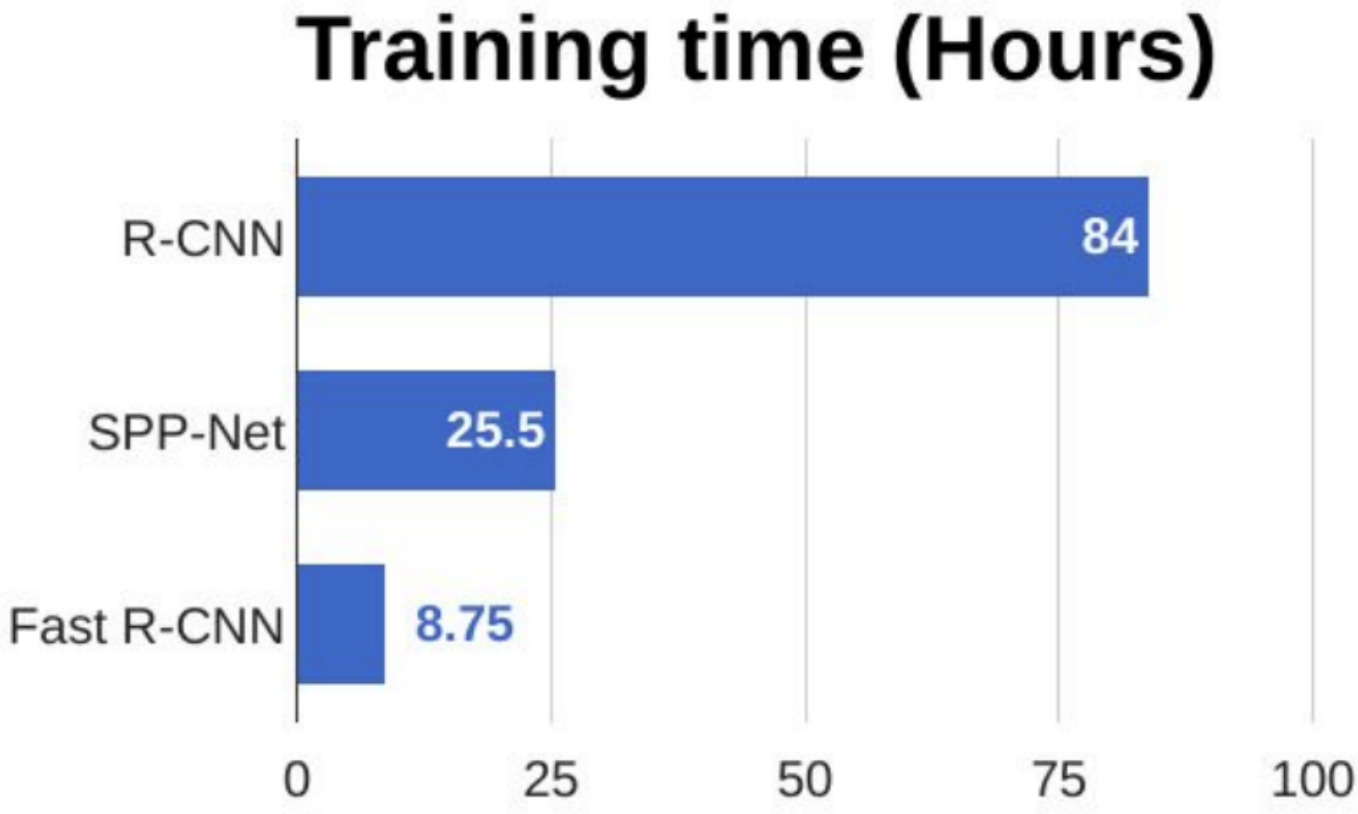


Figure: Training time

Results

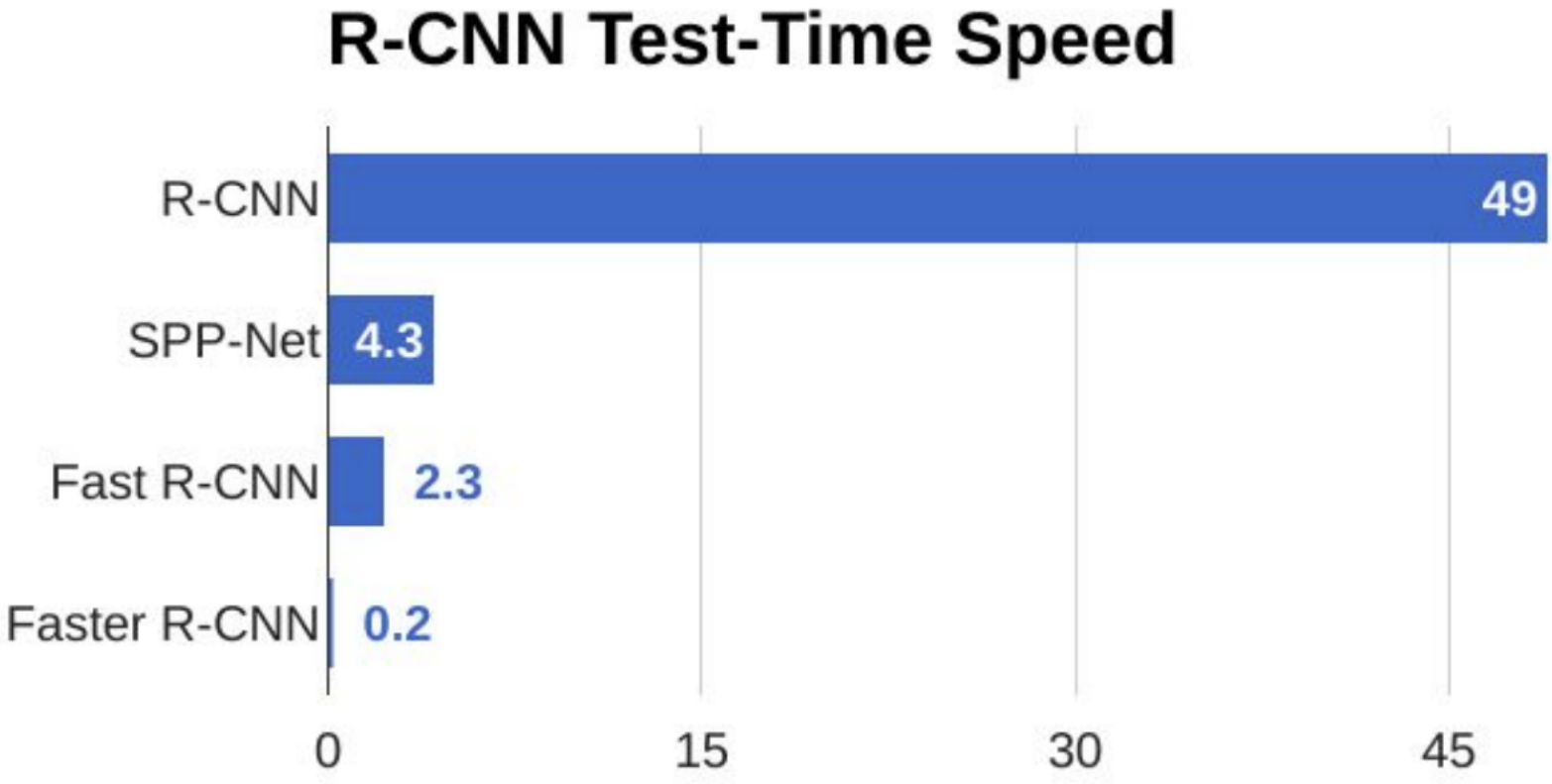


Figure: Test time

Results

Algorithm	Features	Limitations
R-CNN	Selective search for region identification Extraction of CNN features from each region independently for classification	Training is expensive and slow The process involves 3 separate models without much shared computation Extracts around 2000 regions from each image Cannot be implemented in real time (it takes ~47s for each image)
Fast R-CNN	Each image is passed only once to the CNN Feature maps are used to detect objects Uses a single R-CNN model Much faster than R-CNN in both training and testing time	Selective search is slow and hence computation time is high Region proposals are generated separately using a different model. This makes the process very expensive
Faster R-CNN	Uses a unified model composed of RPN (region proposal network) and fast R-CNN with shared convolutional feature layers	Object proposals with RPN are time-consuming The performance of the previous system affects the performance of the current system
Mask R-CNN	Applies Faster R-CNN to pixel-level image segmentation An additional branch is used in parallel with existing branches, to predict an object mask. Improves the RoI pooling layer so that RoI can be more precisely mapped to regions of the original image	None

Figure: Comparative table

Table of contents

Practical session

Figure: Start image

Practical session

Figure: Output image

Practical session

Figure: Output Video

Table of contents

Source

S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *CoRR*, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506.01497

K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, "Mask R-CNN," *CoRR*, vol. abs/1703.06870, 2017. [Online]. Available: http://arxiv.org/abs/1703.06870

R. B. Girshick, "Fast R-CNN," *CoRR*, vol. abs/1504.08083, 2015. [Online]. Available: http://arxiv.org/abs/1504.08083

J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640

V. S. Chandel, "Selective search for object detection," Sep 2017. [Online]. Available: https://www.learnopencv.com/selective-search-for-object-detection-cpp-python/

Source (cont.)

Geva, "Building faster r-cnn on tensorflow: Introduction and examples." [Online]. Available: https://missinglink.ai/guides/tensorflow/building-faster-r-cnn-on-tensorflow-introduction-and-examples/

S. Mallick, "Histogram of oriented gradients," Dec 2016. [Online]. Available: https://www.learnopencv.com/histogram-of-oriented-gradients/

A. SACHAN, "Zero to hero: Guide to object detection using deep learning: Faster r-cnn,yolo,ssd," Dec 2017. [Online]. Available: https://cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd/

P. Sharma, "Step-by-step implementation of mask r-cnn for image segmentation," Jul 2019. [Online]. Available: https://www.analyticsvidhya.com/blog/2019/07/computer-vision-implementing-mask-r-cnn-image-segmentation/