讲师： 贾志刚

OpenCV4 深度神经网络(DNN)实战教程

对象检测模型介绍

- 从分类到检测
- 支持的对象检测网络
- 类别与位置信息解析

从分类到检测

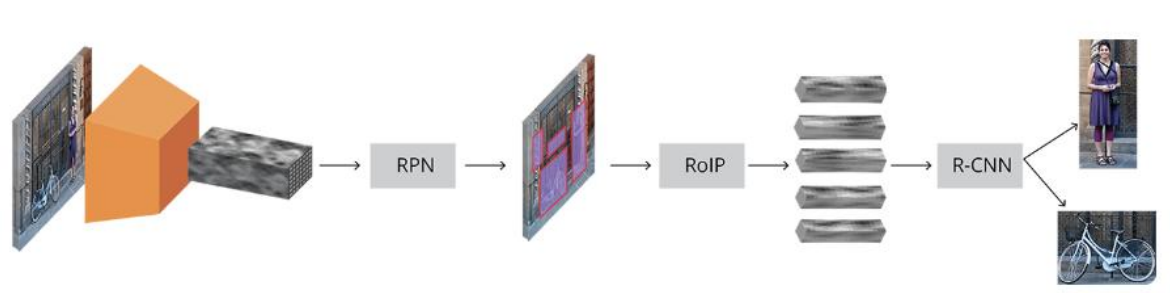
图像分类

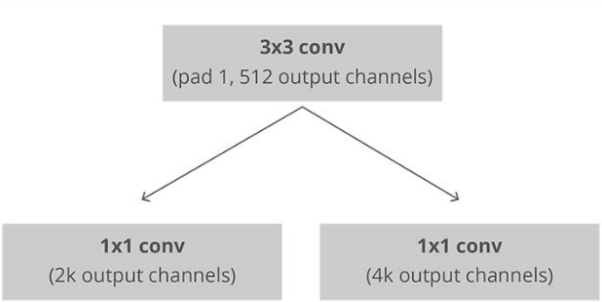
图像分类 + 位置信息 = 对象检测

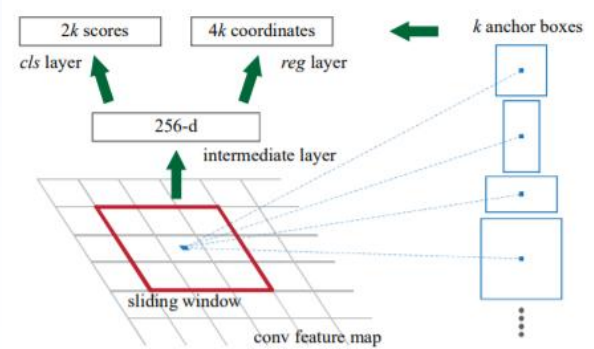
支持的对象检测网络

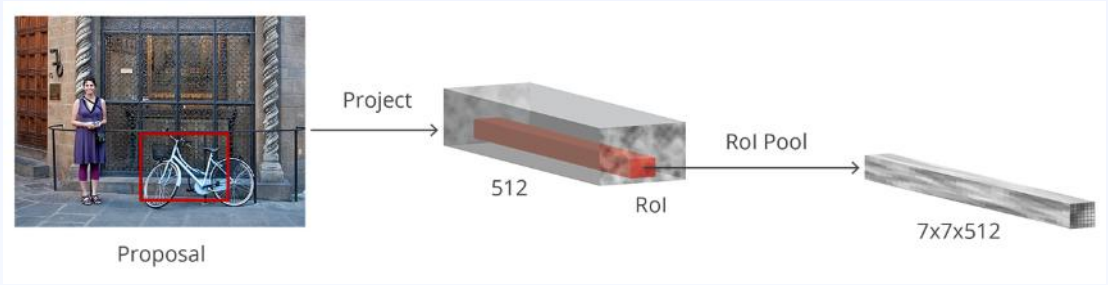
- Faster-RCNN
- SSD(VGG/mobile-net backbone)
- YOLO(YOLOv3/YOLOv3-tiny)

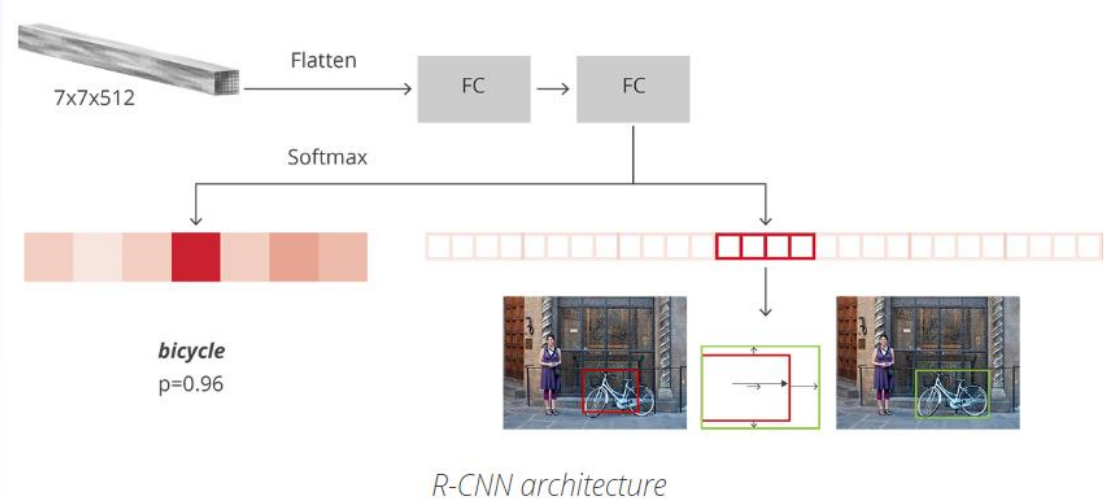
Faster-RCNN


RPN
RoIP
R-CNN


3x3 conv
(pad 1, 512 output channels)
1x1 conv
(2k output channels)
1x1 conv
(4k output channels)

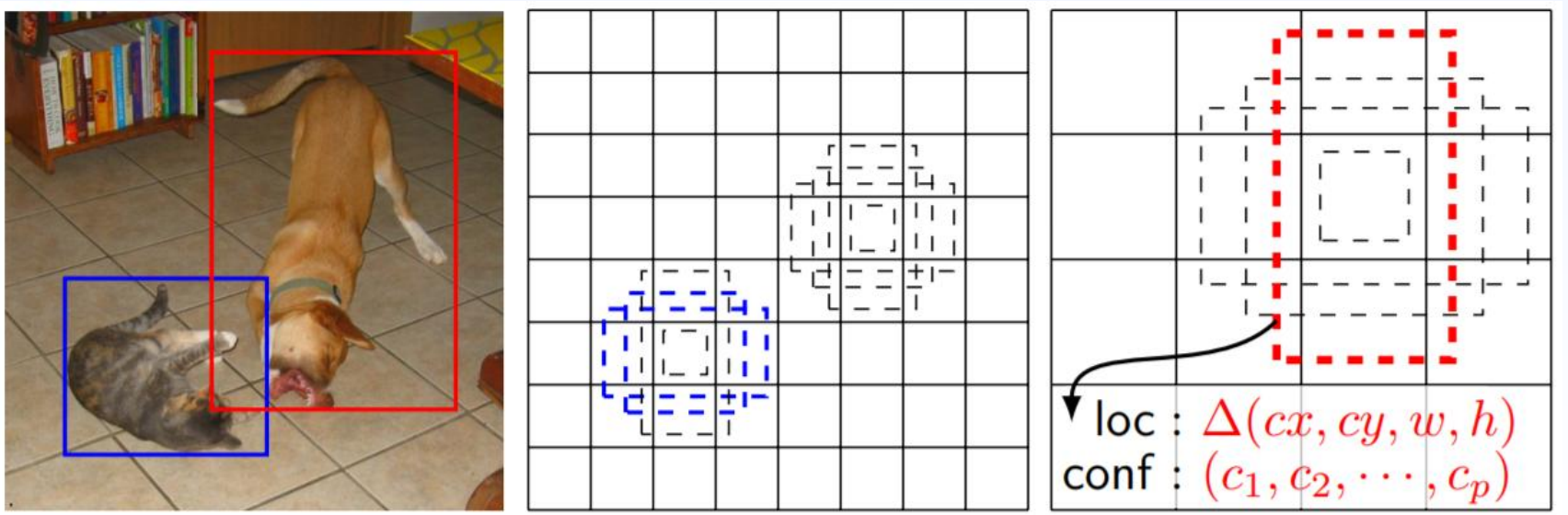

2k scores
4k coordinates
k anchor boxes
cls layer
reg layer
256-d
intermediate layer
sliding window
conv feature map





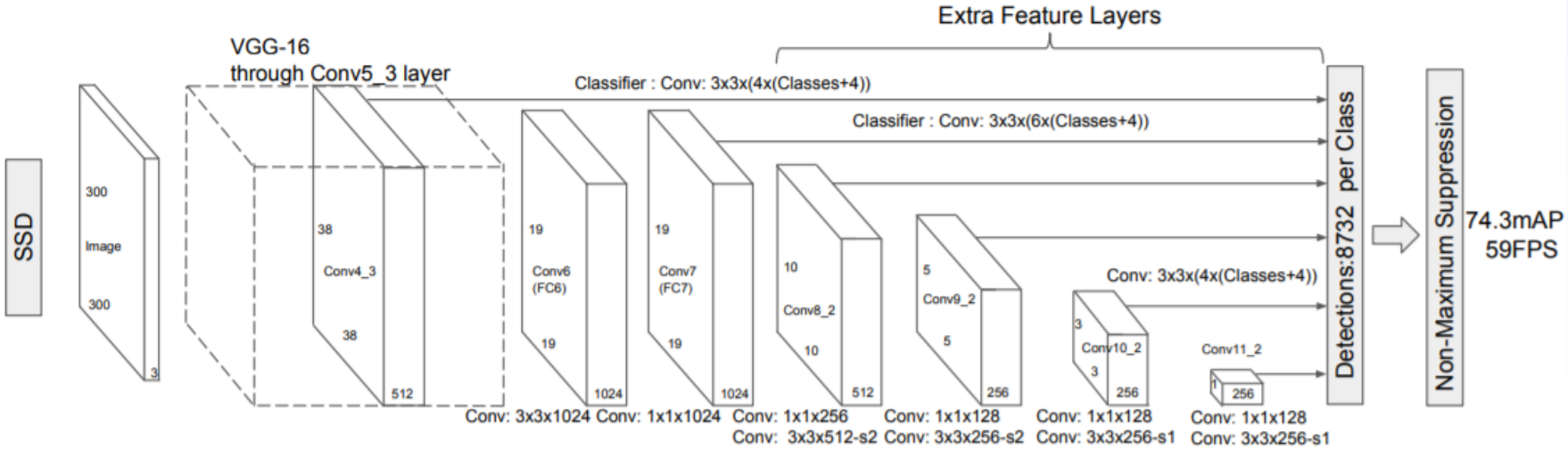
R-CNN architecture

SSD

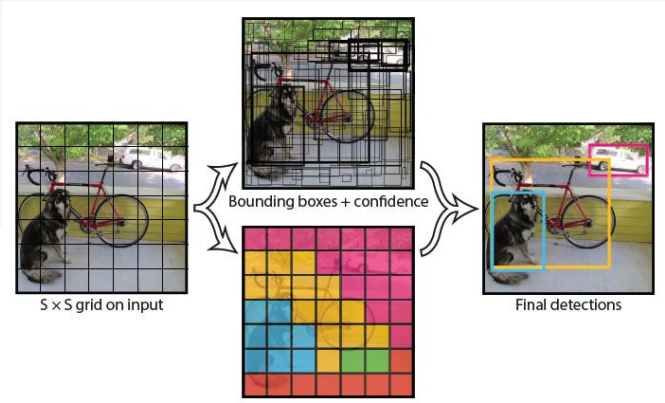


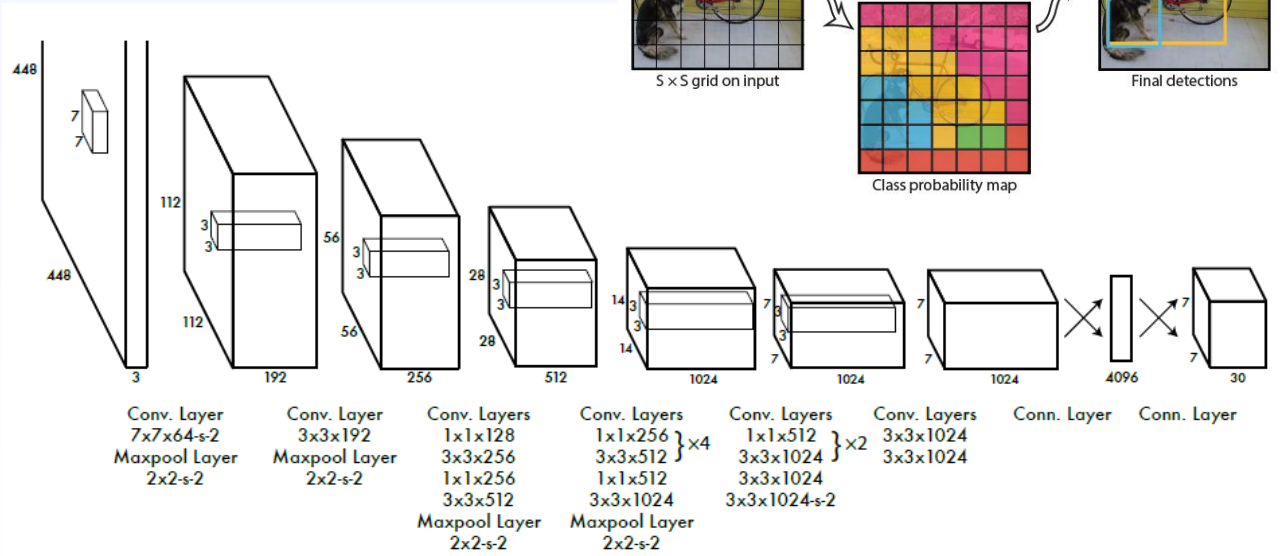
(a) Image with GT boxes (b) 8×8 feature map (c) 4×4 feature map

SSD


SSD
300
Image
300
3
VGG-16
through Conv5_3 layer
38
Conv4_3
38
512
19
Conv6
(FC6)
19
1024
19
Conv7
(FC7)
19
1024
10
Conv8_2
10
512
5
Conv9_2
5
256
3
Conv10_2
3
256
Conv11_2
1
256
Extra Feature Layers
Classifier : Conv: 3x3x(4x(Classes+4))
Classifier : Conv: 3x3x(6x(Classes+4))
Conv: 3x3x(4x(Classes+4))
Detections:8732 per Class
Non-Maximum Suppression
74.3mAP
59FPS
Conv: 3x3x1024
Conv: 1x1x1024
Conv: 1x1x256
Conv: 3x3x512-s2
Conv: 1x1x128
Conv: 3x3x256-s2
Conv: 1x1x128
Conv: 3x3x256-s1
Conv: 1x1x128
Conv: 3x3x256-s1

YOLO-v1


Bounding boxes + confidence
S × S grid on input
Final detections
Class probability map

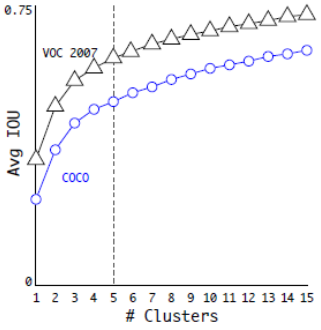

448
448
7
7
3
112
112
192
3
3
56
56
256
3
3
28
28
512
3
3
14
14
1024
3
3
7
7
1024
3
3
7
7
1024
4096
7
7
30
Conv. Layer
7x7x64-s-2
Maxpool Layer
2x2-s-2
Conv. Layer
3x3x192
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x128
3x3x256
1x1x256
3x3x512
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x256
3x3x512
} ×4
1x1x512
3x3x1024
Maxpool Layer
2x2-s-2
Conv. Layers
1x1x512
3x3x1024
} ×2
3x3x1024-s-2
Conv. Layers
3x3x1024
3x3x1024
Conn. Layer
Conn. Layer

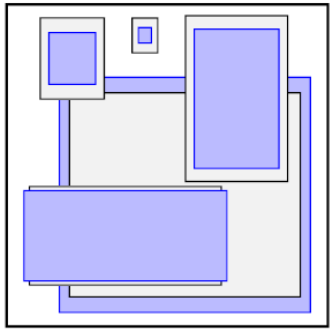
YOLO-v3

- 多尺度输出
- NxNx[3*(4+1+80)]
- 输出cx, cy, h, w

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Table 1. **Darknet-53.**





类别与位置信息解析

- Detection out/info层 - [1x1xNx7]
- Region层 - [1x1xNx(5+score)], 取决于数据集

Thank You!