U11. CNNs in Computer Vision

SJK002 Computer Vision

Master in Intelligent Systems

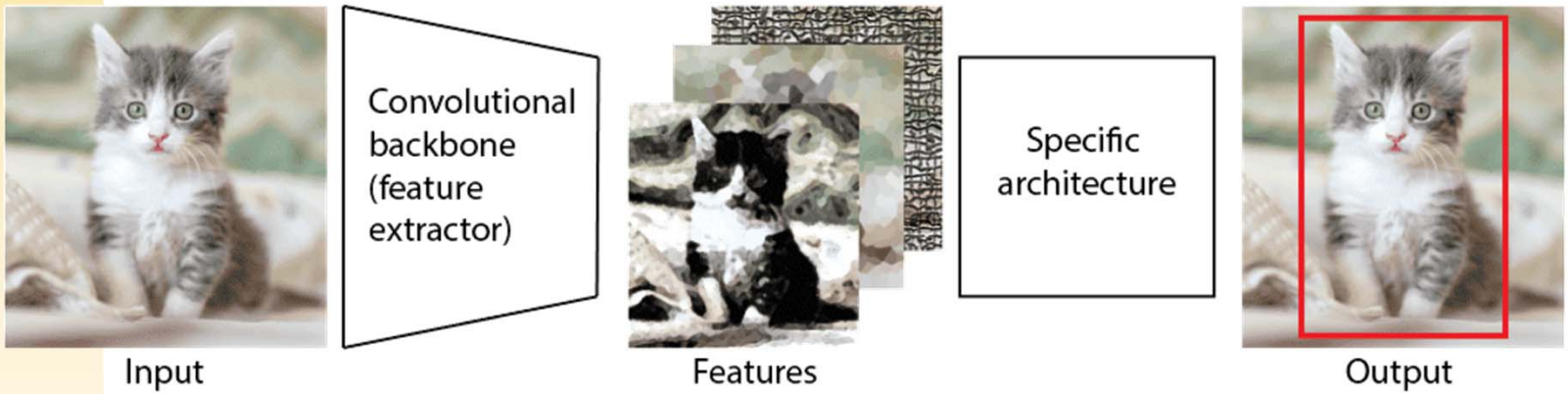
UNIVERSITAT JAUME·I

UNIVERSITAT JAUME I

Index

- Image classification: Vanilla CNNs
- Backbones:
  - VCC, ResNet, …
- Object detection (YOLO)
- U-Net:
  - Image segmentation
  - Optical flow estimation
- Long Short Term Memory (LSTM):
  - Temporal series
- Transformers:
  - Image classification
  - Image segmentation

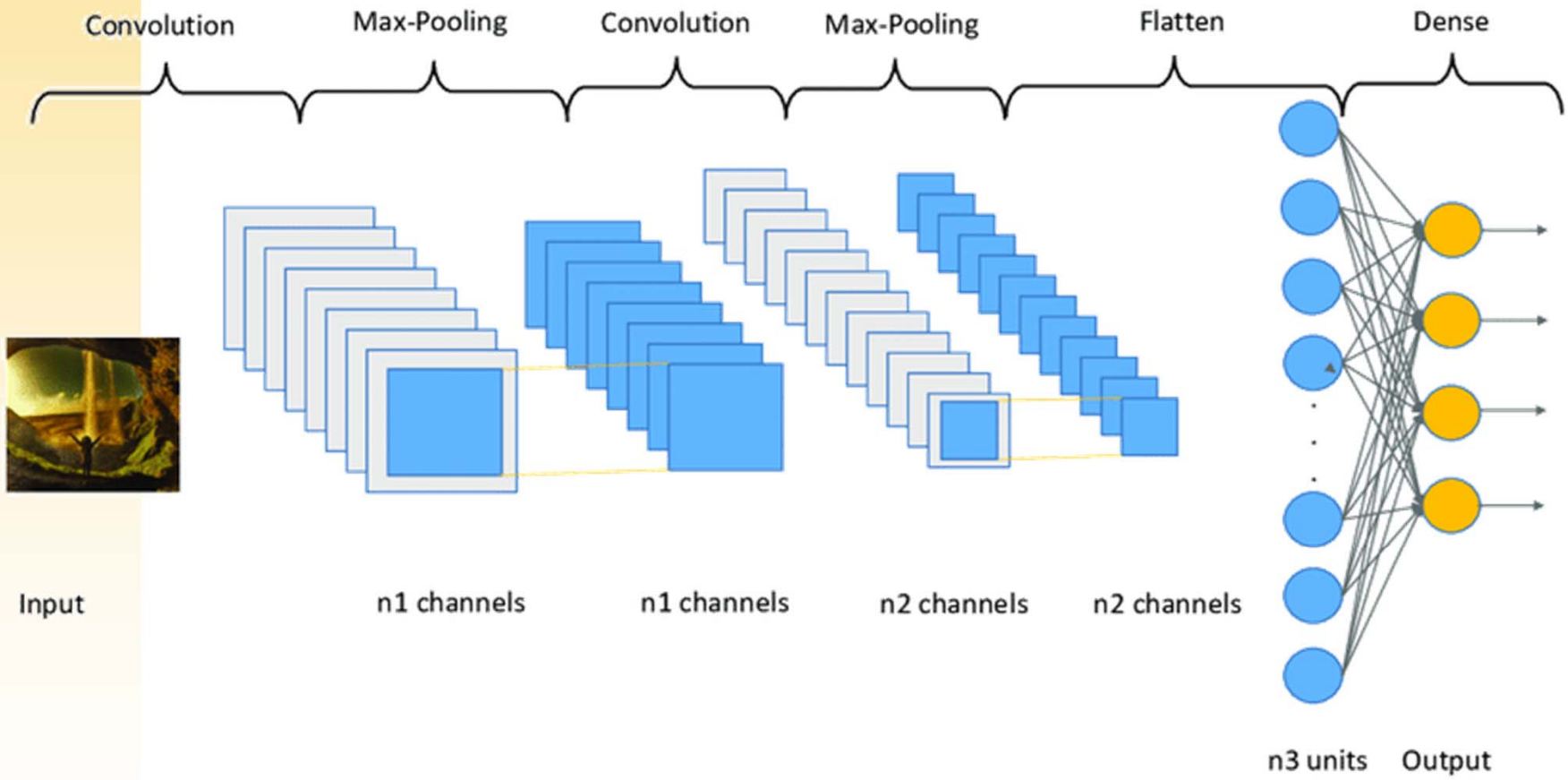
CNN models for CV



Feature learning/extraction

Classification
Segmentation
Regression
Detection
Generation
…

Image classification: Vanilla CNNs



Vanilla CNNs

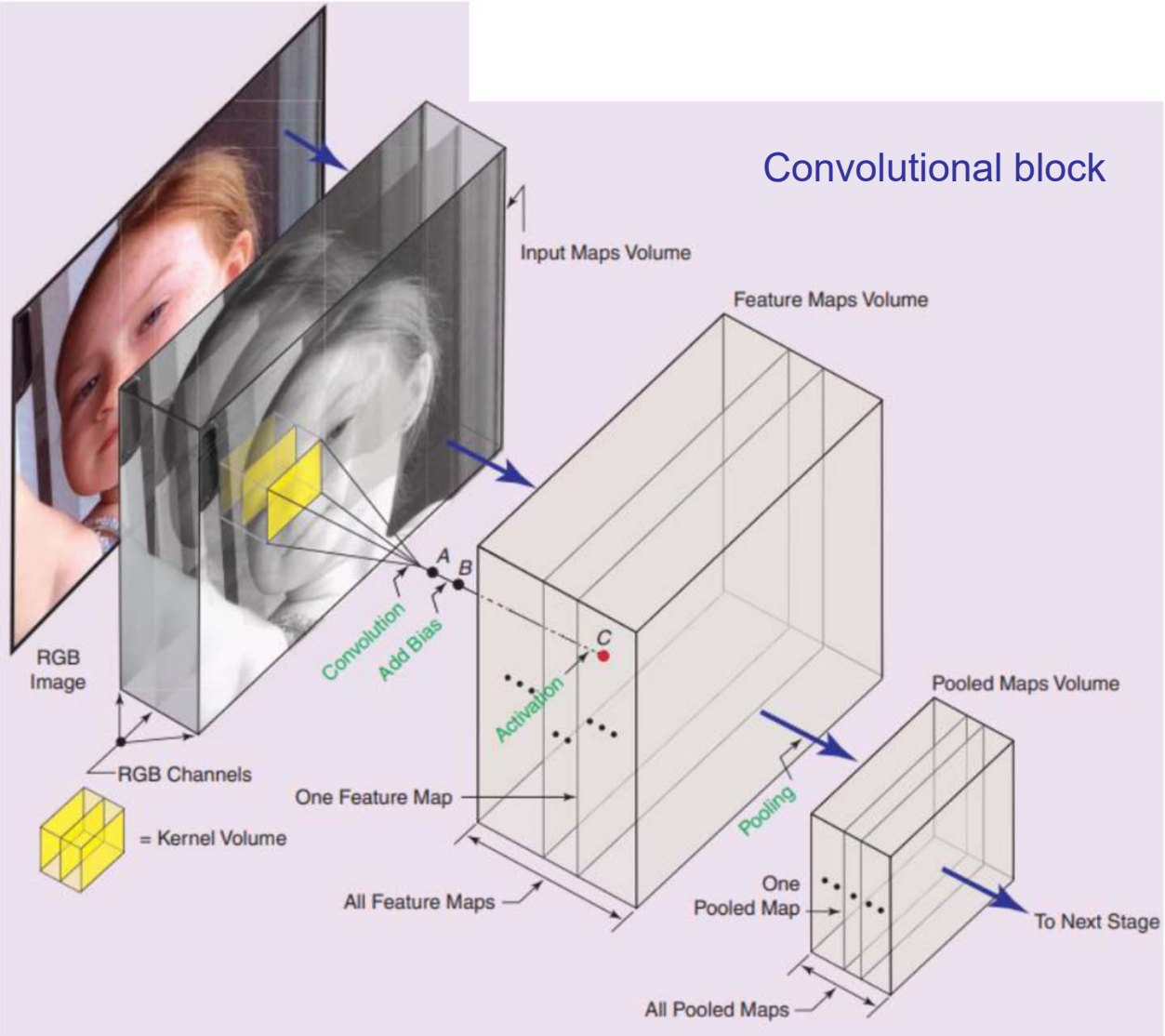
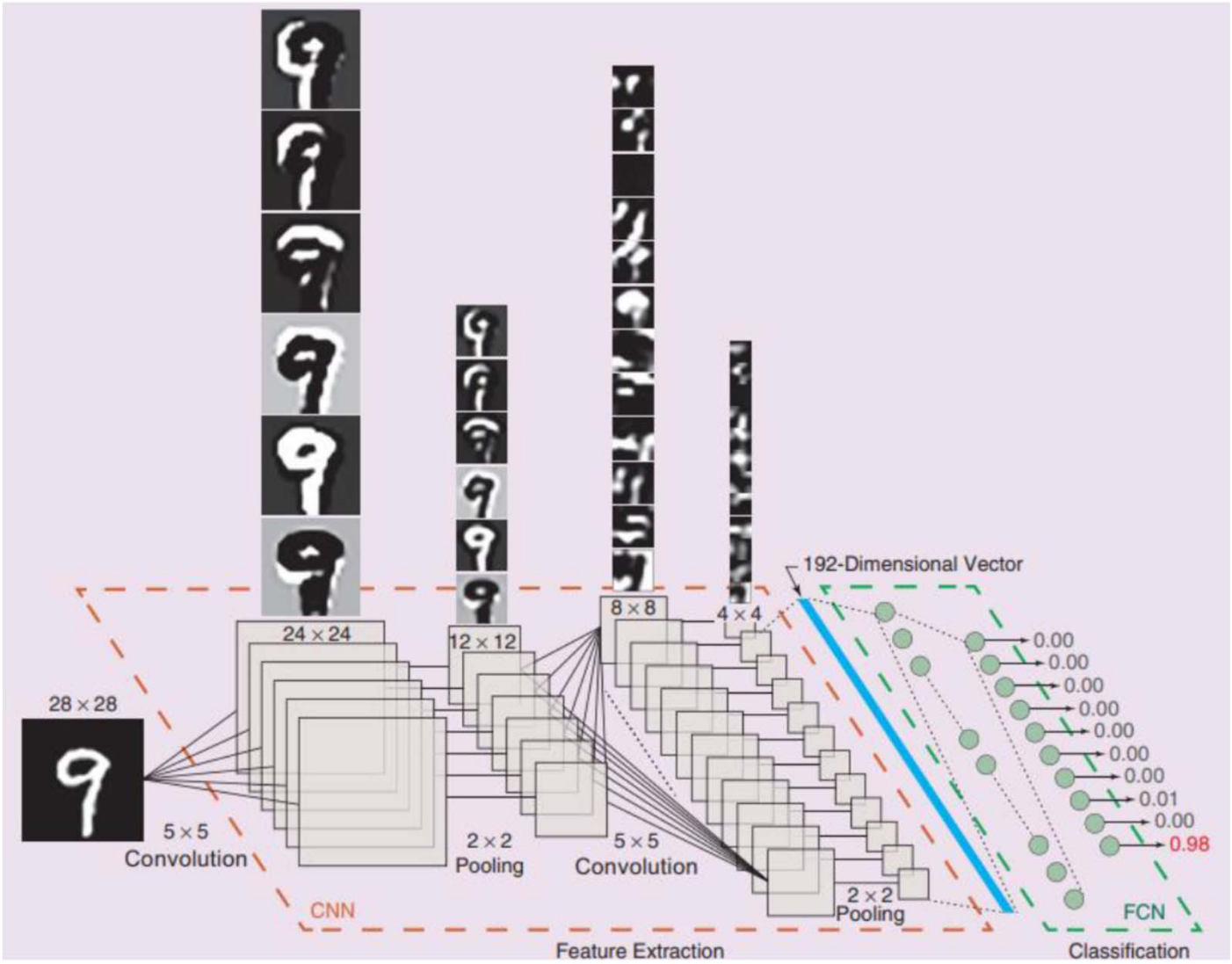
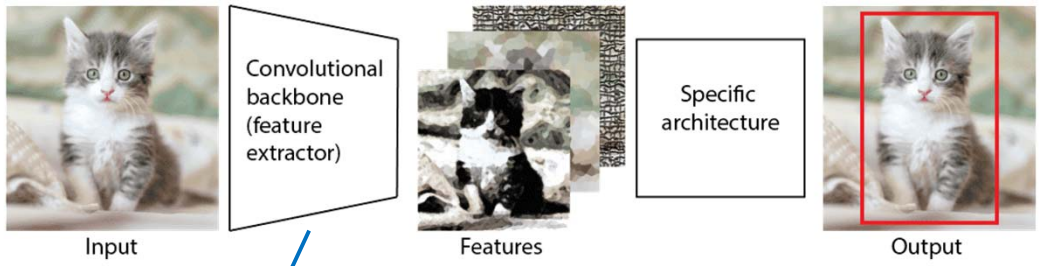
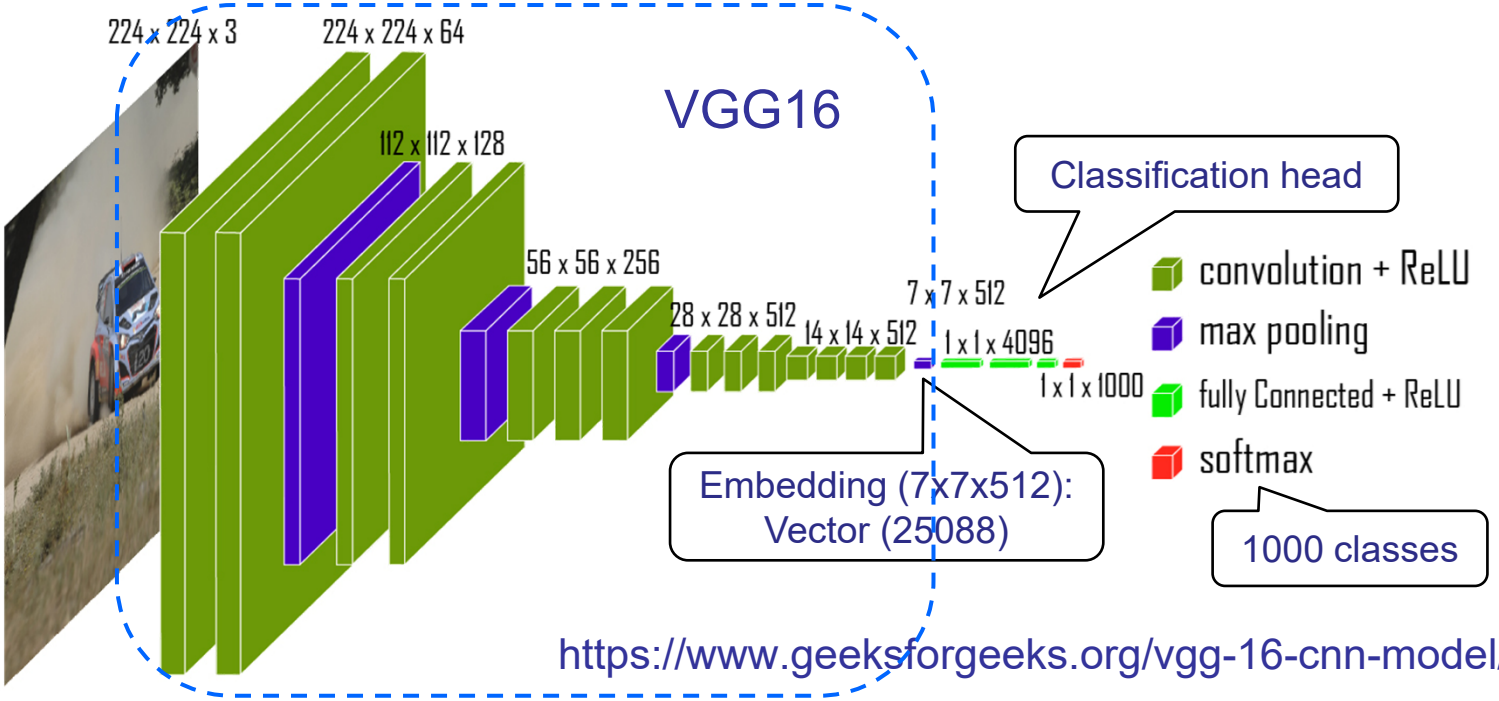


Image classification: Vanilla CNNs



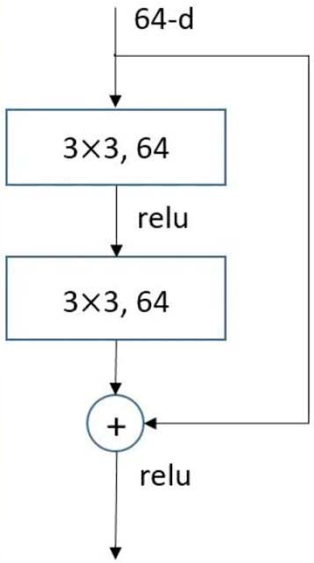
Pre-trained backbones: VGG16



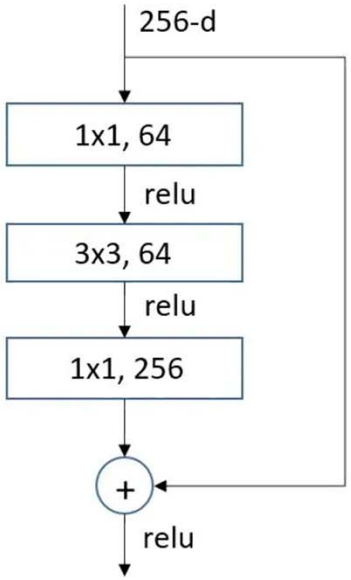


https://www.geeksforgeeks.org/vgg-16-cnn-model/

Pre-trained backbones: ResNet



Residual block used in ResNet 18, ResNet 34

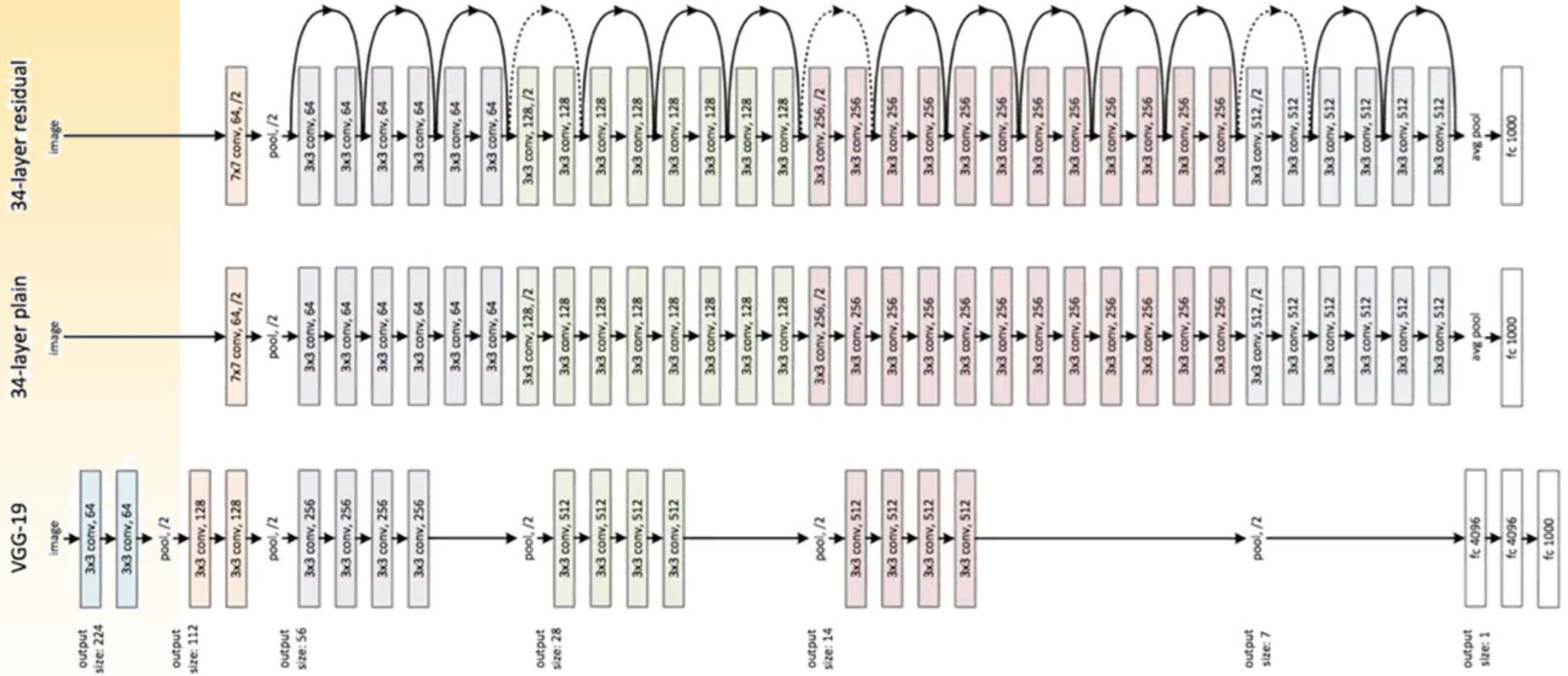


Residual block with bottleneck used in ResNet 50, 101, 152

params of bottleneck
1x1x256x64+3x3x64x64+1x1x64x256=69632

params of two 3x3 convolutions
3x3x256x64+3x3x64x256=294912

Pre-trained backbones: ResNet



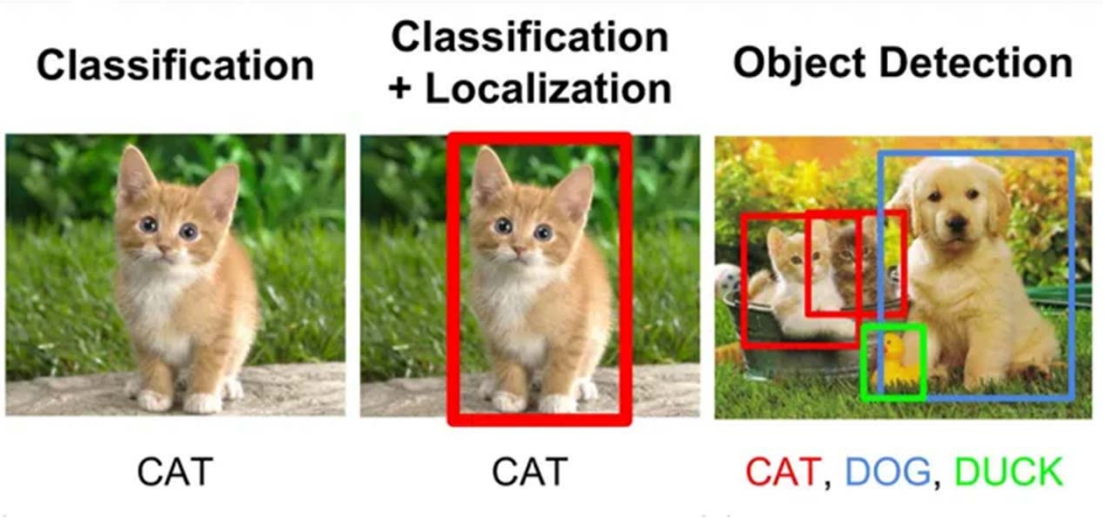
UNIVERSITAT JAUME I

Pre-trained backbones

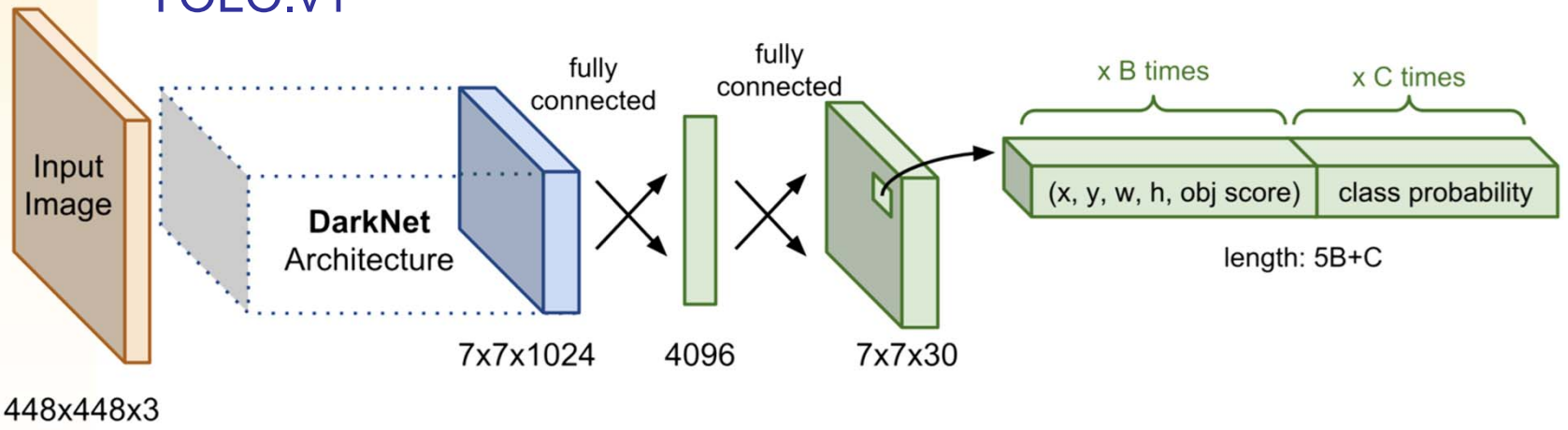
Backbone	Year	# of parameters	trained task
AlexNet [6]	2012	60M	Img-class
VGG-16 [10]	2014	138M	Img-class
VGG-19 [10]	2014	144M	Img-class
Inception-V1 (GoogleNet) [16]	2014	5 M	Img-class
ResNet-18 [17]	2015	11.7 M	Img-class
ResNet-34 [17]	2015	25.6 M	Img-class
ResNet-50 [17]	2015	26 M	Img-class
ResNet-101 [17]	2015	44.6 M	Img-class
ResNet-152 [17]	2015	230M	Img-class
Inception-v2 [20]	2015	21.8M	Img-class
Inception-v3 [20]	2015	21.8M	Img-class
Inception-ResNet-V2 [21]	2015	55 M	Img-class, obj-det
Darknet-19 2015 [28]	2015	20.8 M	Obj-det
Xception [41]	2017	22.9 M	Img-class
SqueezeNet 2016 [36]	2016	1.24M	Img-class
ShuffleNet[42](g = 1)	2017	143M M	Img-class, obj-det
ShuffleNet-v2[43](g = 1)	2018	2.3 M	Img-class, obj-det
DenseNet-100 (k = 12) [18]	2018	7.0M	Img-class
DenseNet-100 (k = 24) [18]	2018	27.2M	Img-class
DenseNet-250 (k = 24) [18]	2018	15.3M	Img-class
DenseNet-190 (k = 40) [18]	2018	25.6M	Img-class
DetNet [35]	2018	-	Img-class, obj-det
EfficientNet B0 to B7 [44]	2020	5.3 M, to 66M	Img-class, obj-det
MobileNet [38]	2017	4.2 M	Img-class, obj-det
MobileNet-v2 [39]	2017	3.4 M	Img-class, obj-det
WideResNet-40-4 [37]	2016	8.9 M	Img-class, obj-det
WideResNet-16-8 [37]	2016	11 M	Img-class, obj-det
WideResNet-28-10 [37]	2016	36.5 M	Img-class, obj-det
SWideRNet (w_1=0.25, w_2=0.25, l=0.75) [47]	2020	7.77 M	Panoptic-seg
SWideRNet (w_1=1, w_2=1, l=1) [47]	2020	168.77 M	Panoptic-seg
SWideRNet (w_1=1, w_2=1, l=6) [47]	2020	836.59 M	Panoptic-seg
SWideRNet (w_1=1, w_2=1.5, l=3) [47]	2020	946.69 M	Panoptic-seg
HRNet W32, W48 [45]	2019	28.5M, 63.6M	Human-Pose- est
HRNet V2 [45]	2020	-	Semantic-seg

Backbones review:
https://arxiv.org/pdf/2206.08016

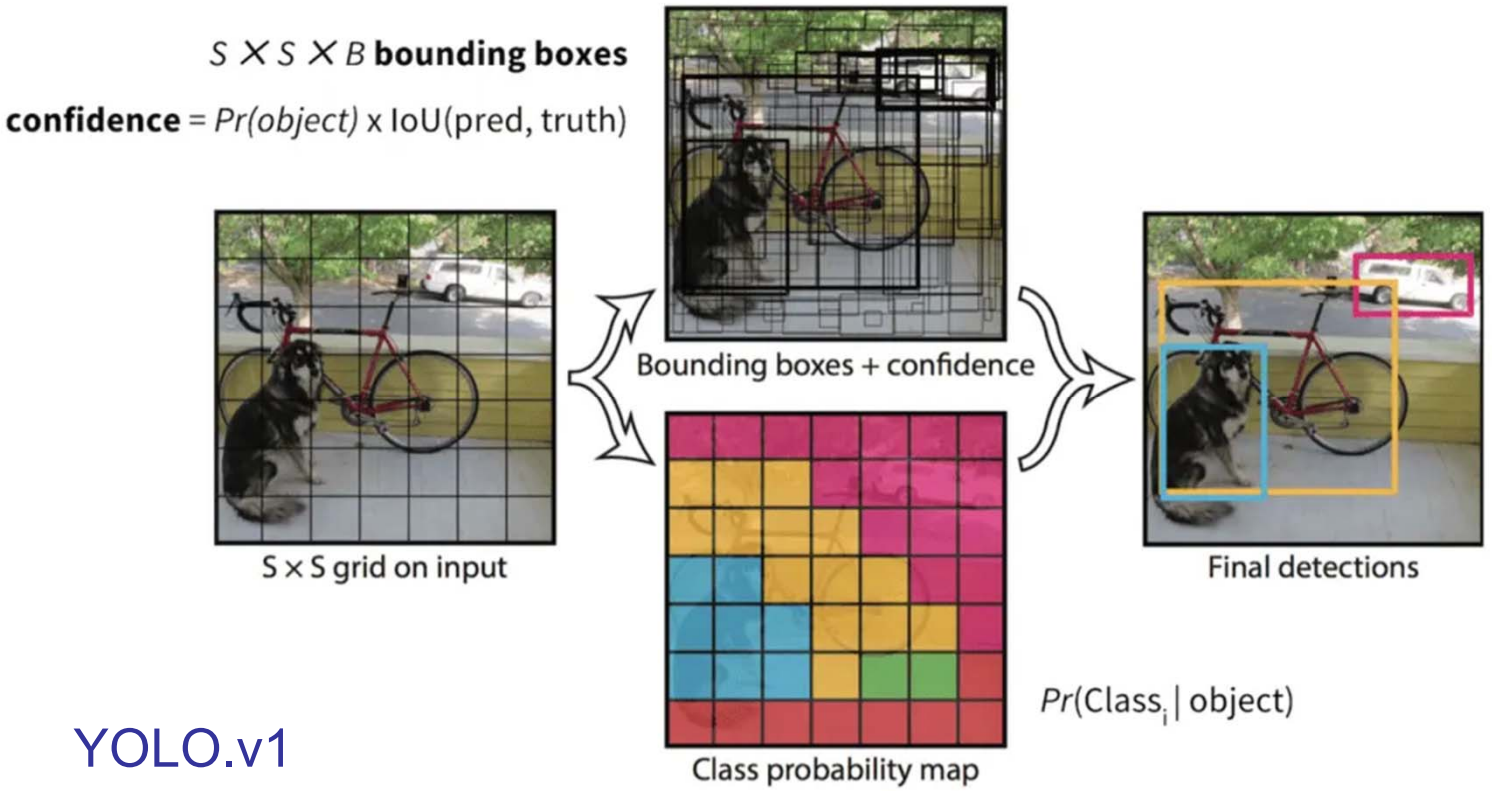
Object detection: YOLO models



YOLO.v1



Object detection: YOLO models



YOLO.v1

https://learnopencv.com/mastering-all-yolo-models/