EVD 3

ADVANCED CNN

JEROEN VEEN

HAN_UNIVERSITY
OF APPLIED SCIENCES

DL REPORT TEMPLATE

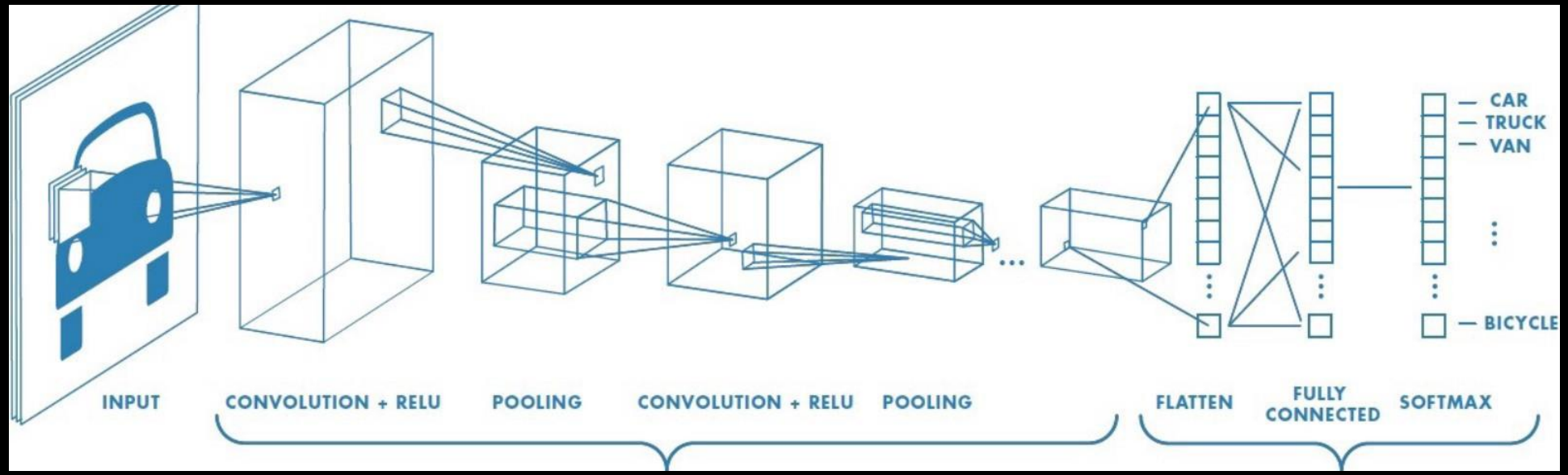
- Questions?

AGENDA

- Object detection
- Object tracking
- Semantic segmentation
- Variational autoencoder

TYPICAL CNN ARCHITECTURE

- Perform classification
- Conv, pool, dense layers



Source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

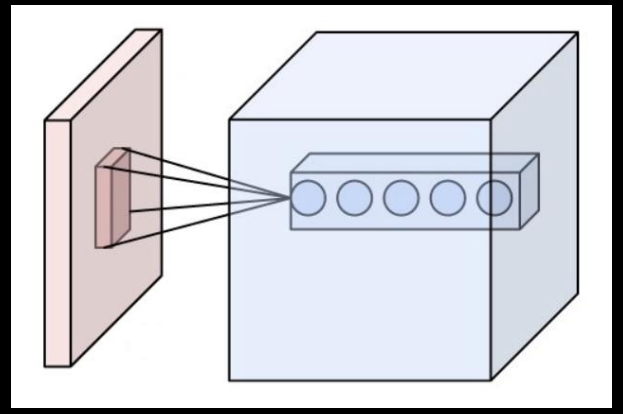
CNN SUMMARY

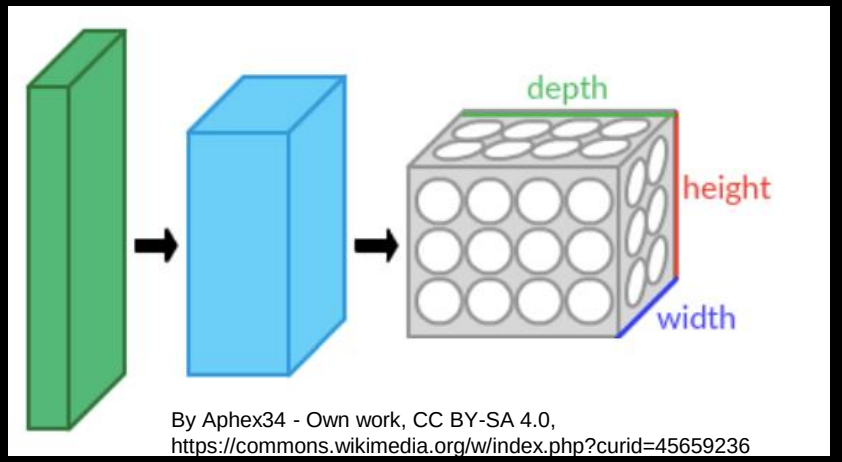
- Emulate the behavior of a visual cortex (e.g. receptive fields)
- Higher-level representations of image content
- No feature definition, but automated extraction

- Biologically inspired perceptrons
- Multilayer perceptrons usually mean fully connected networks, which makes them prone to overfitting
- CNNs can be considered as regularized versions of multilayer perceptrons

CONVOLUTIONAL LAYER SUMMARY

- Local connectivity
- Shared weights
- 3D volumes of neurons
- Output is a stack if feature maps





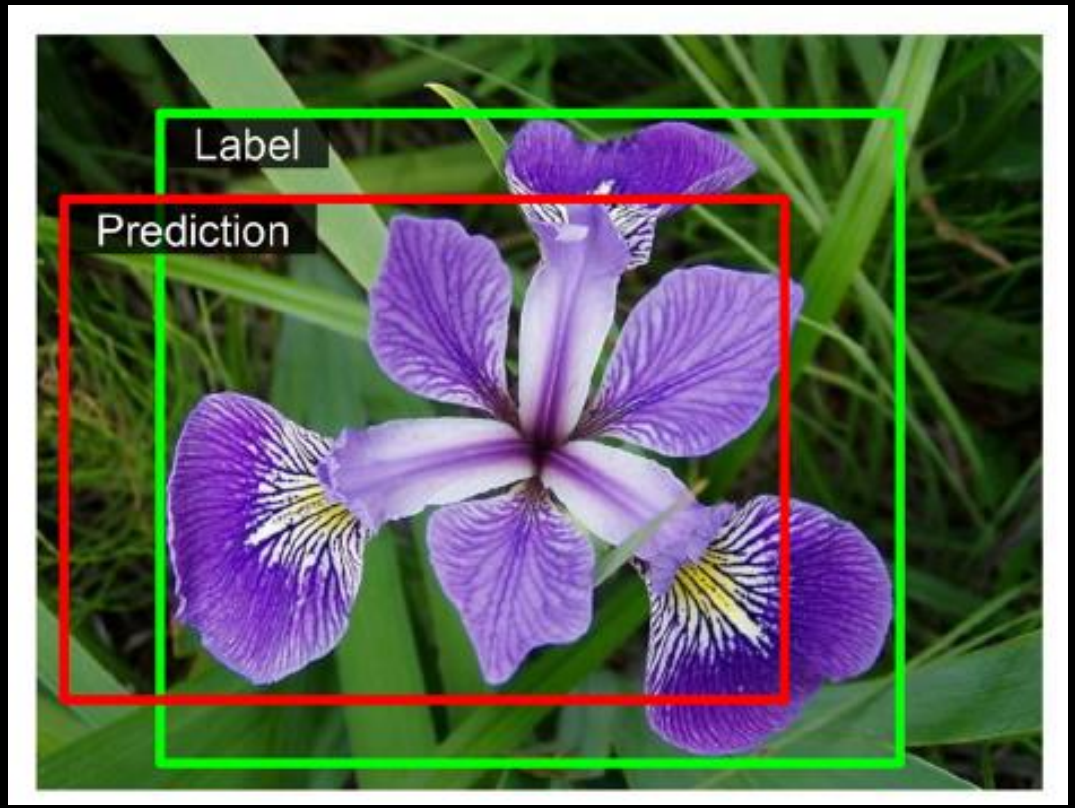
By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45659236

MORE TRANSFER LEARNING HINTS

- https://keras.io/api/applications

CLASSIFICATION AND LOCALIZATION

- Add second dense output layer to predict coordinates (regression)

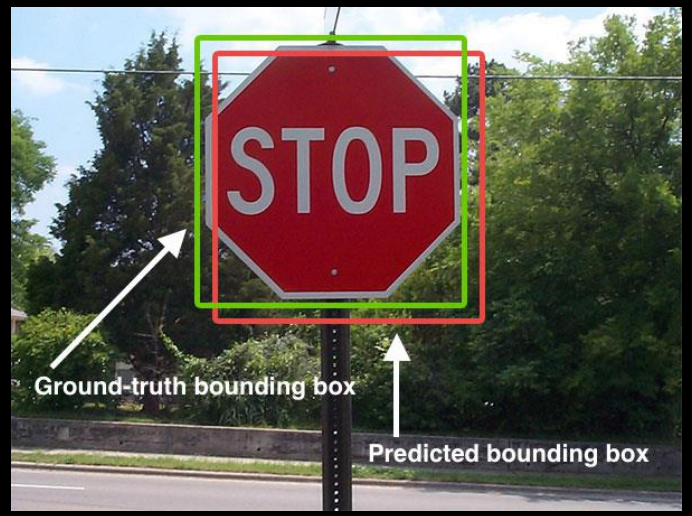


Source: Géron, ISBN: 9781492032632

JACCARD INDEX

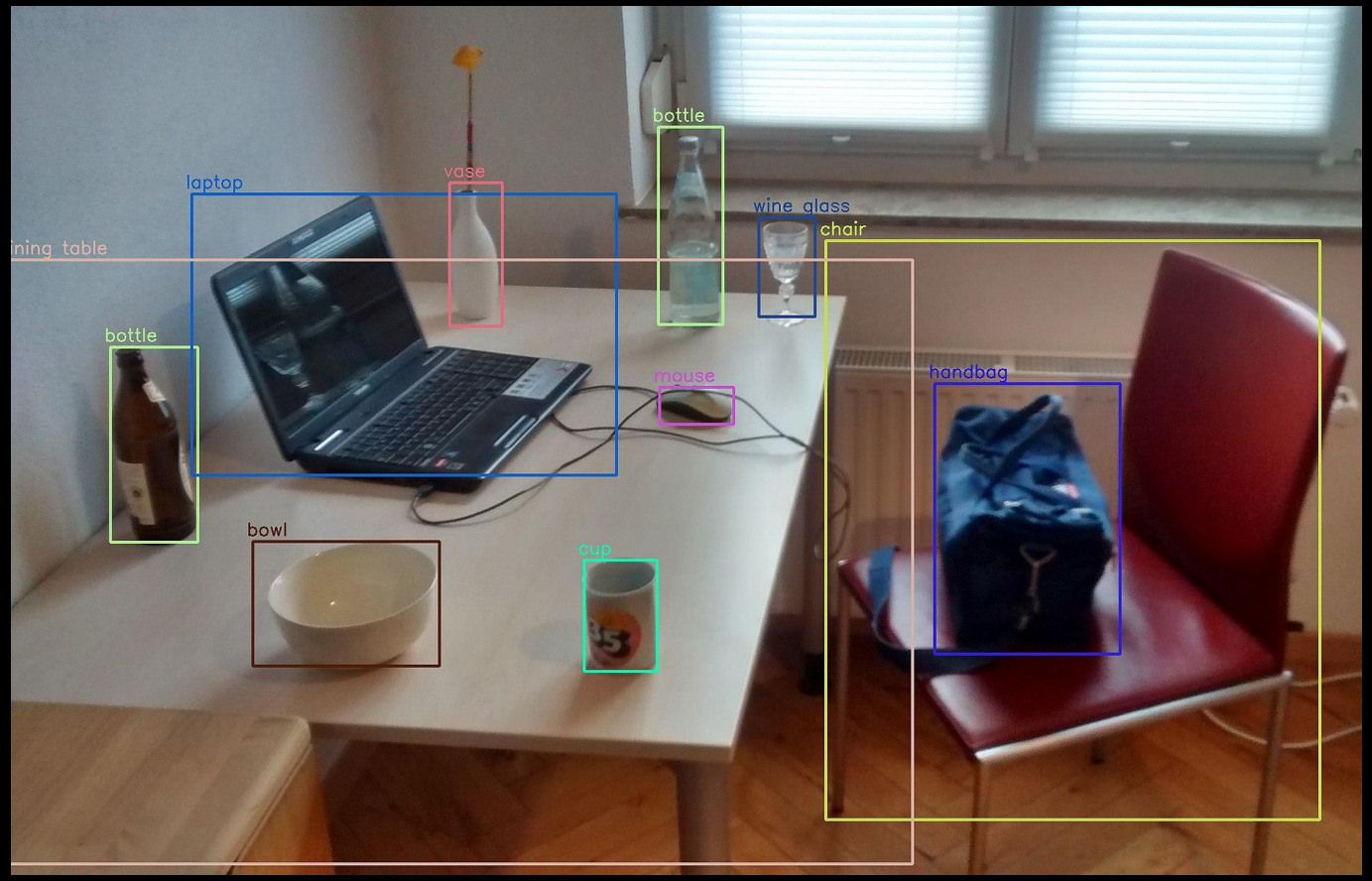
- Performance metric: intersection over Union (IoU)
- measures similarity between finite sample sets

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

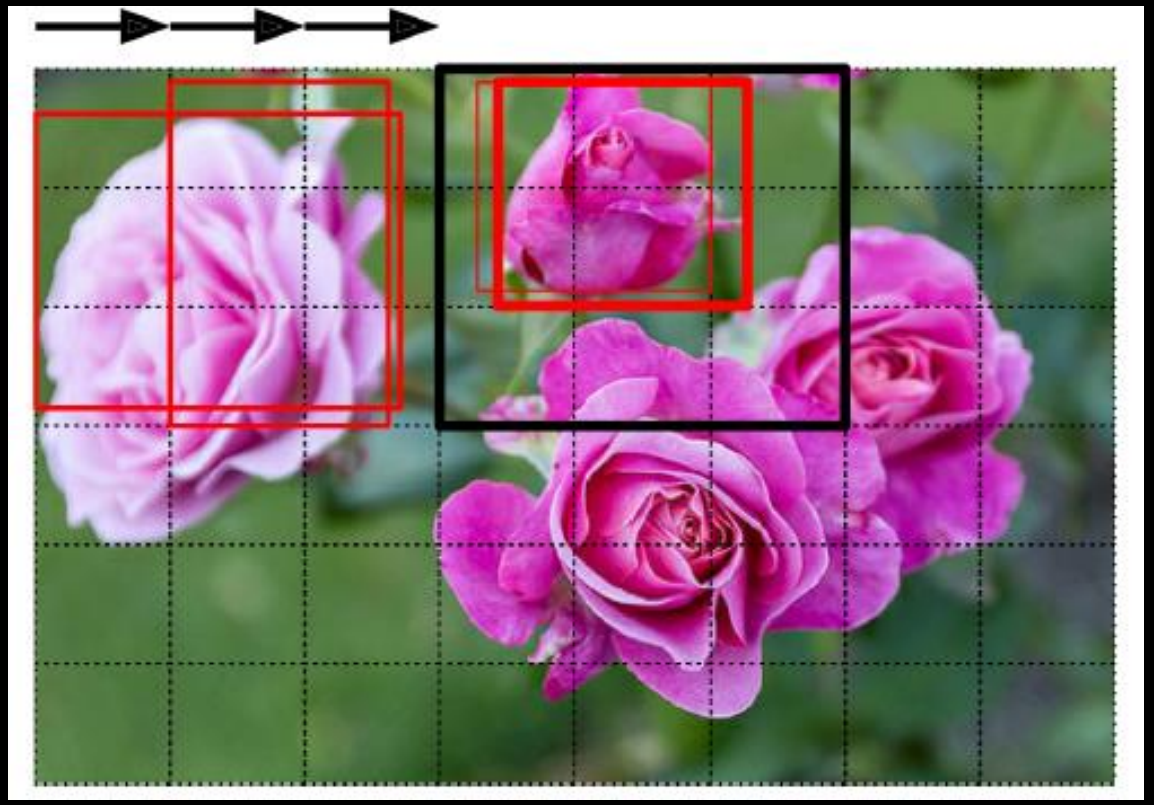


Source: By Adrian Rosebrock - http://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=57718560

OBJECT DETECTION

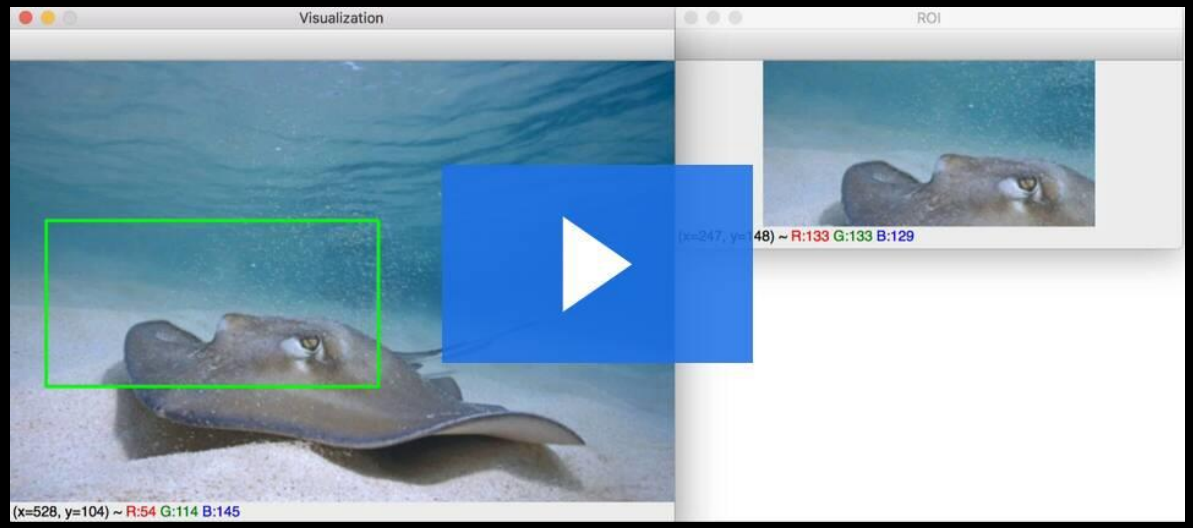


SLIDING WINDOW



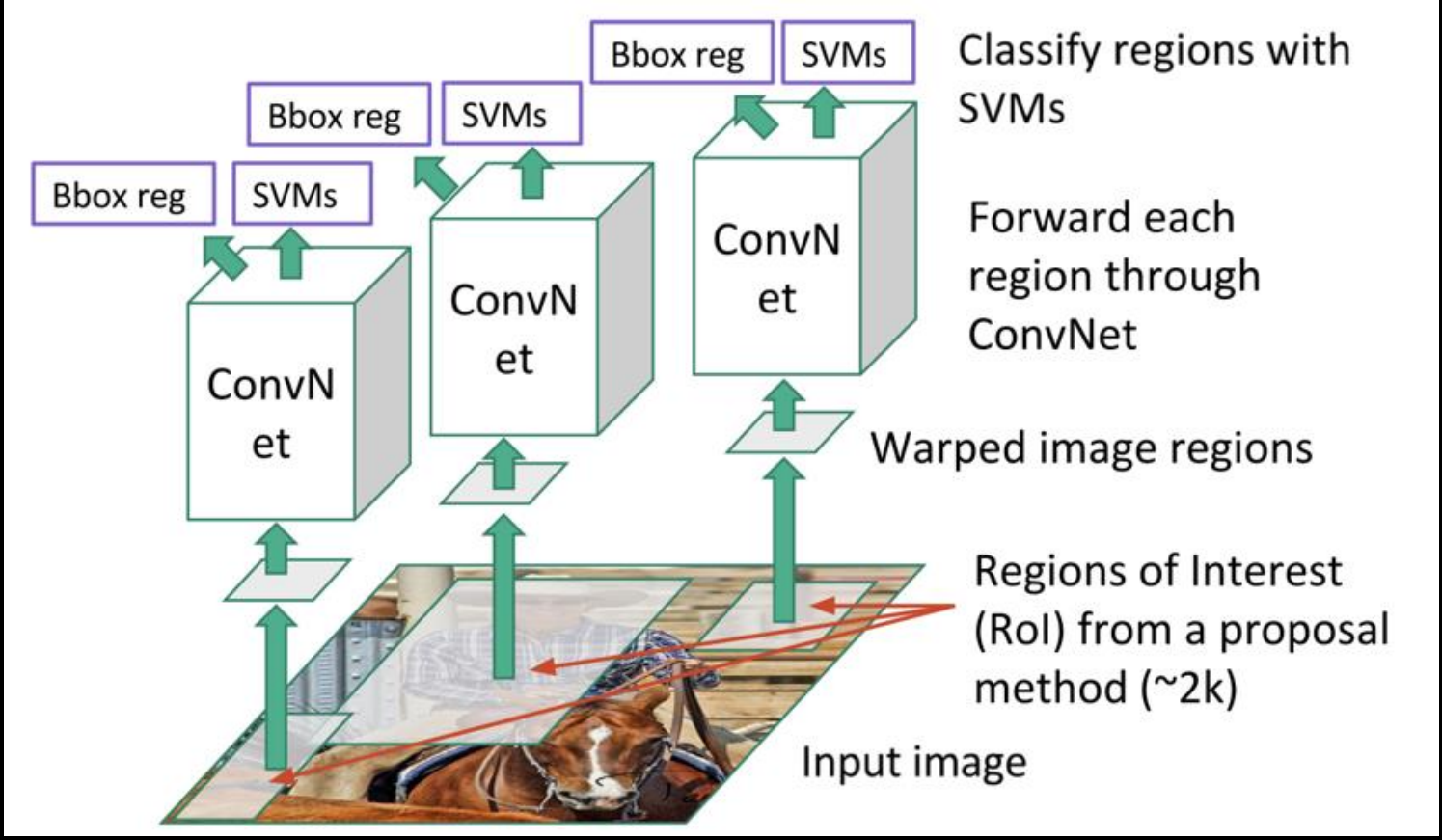
Source: Géron, ISBN: 9781492032632

SLIDING WINDOWS

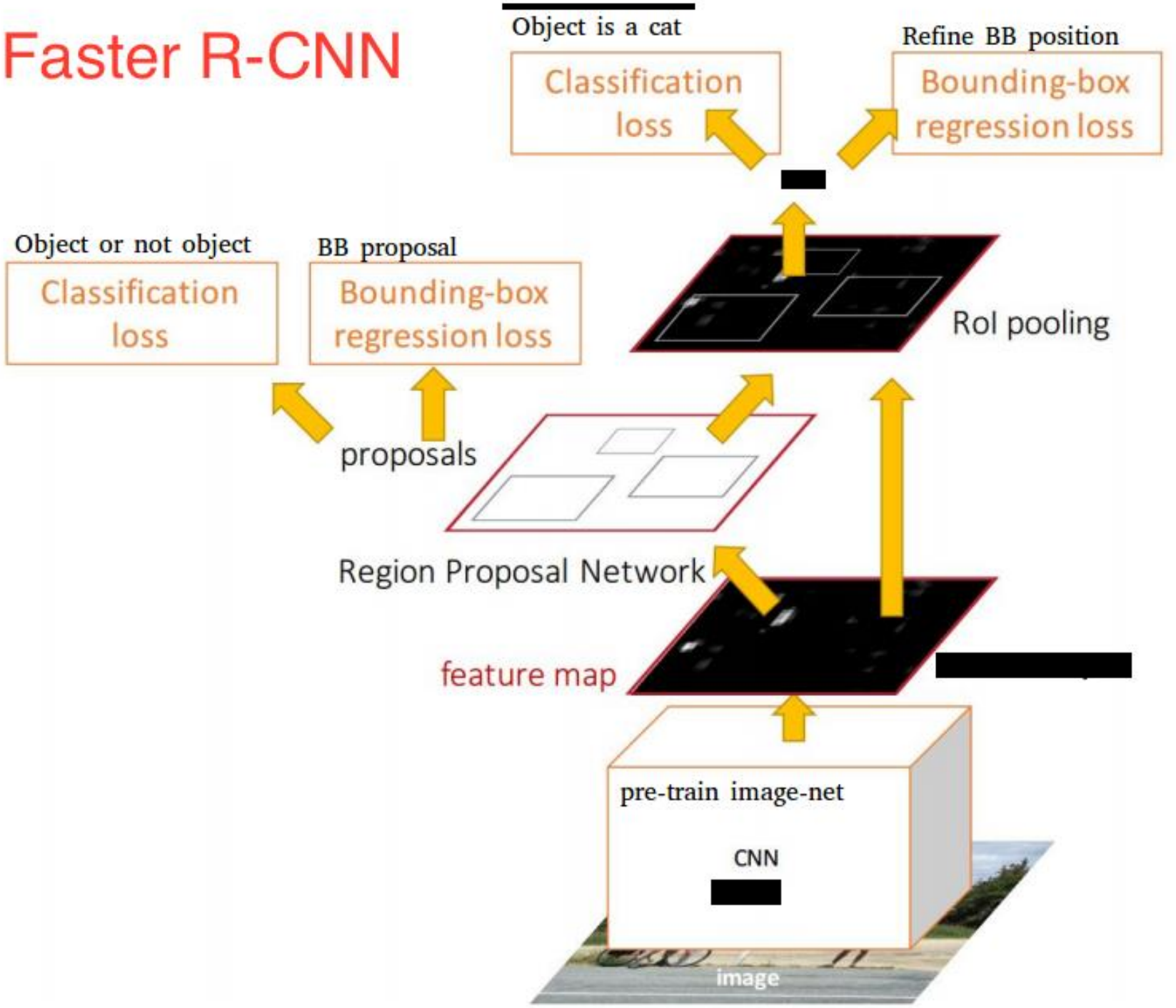


Source: https://www.pyimagesearch.com/2020/06/22/turning-any-cnn-image-classifier-into-an-object-detector-with-keras-tensorflow-and-opencv/

REGION-BASED CNN

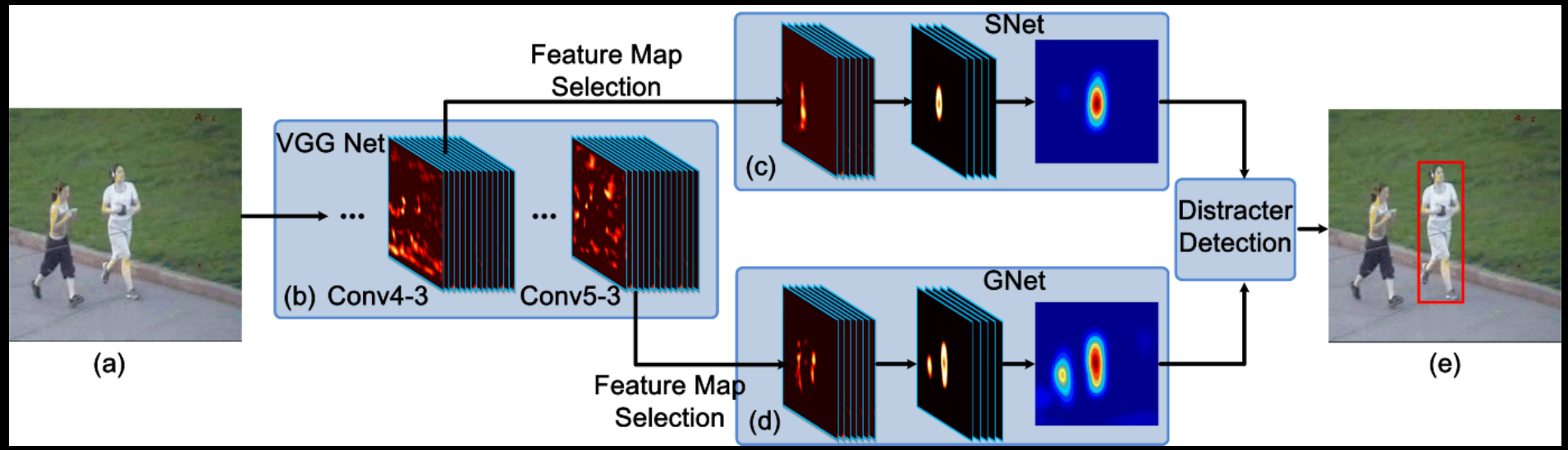


Faster R-CNN

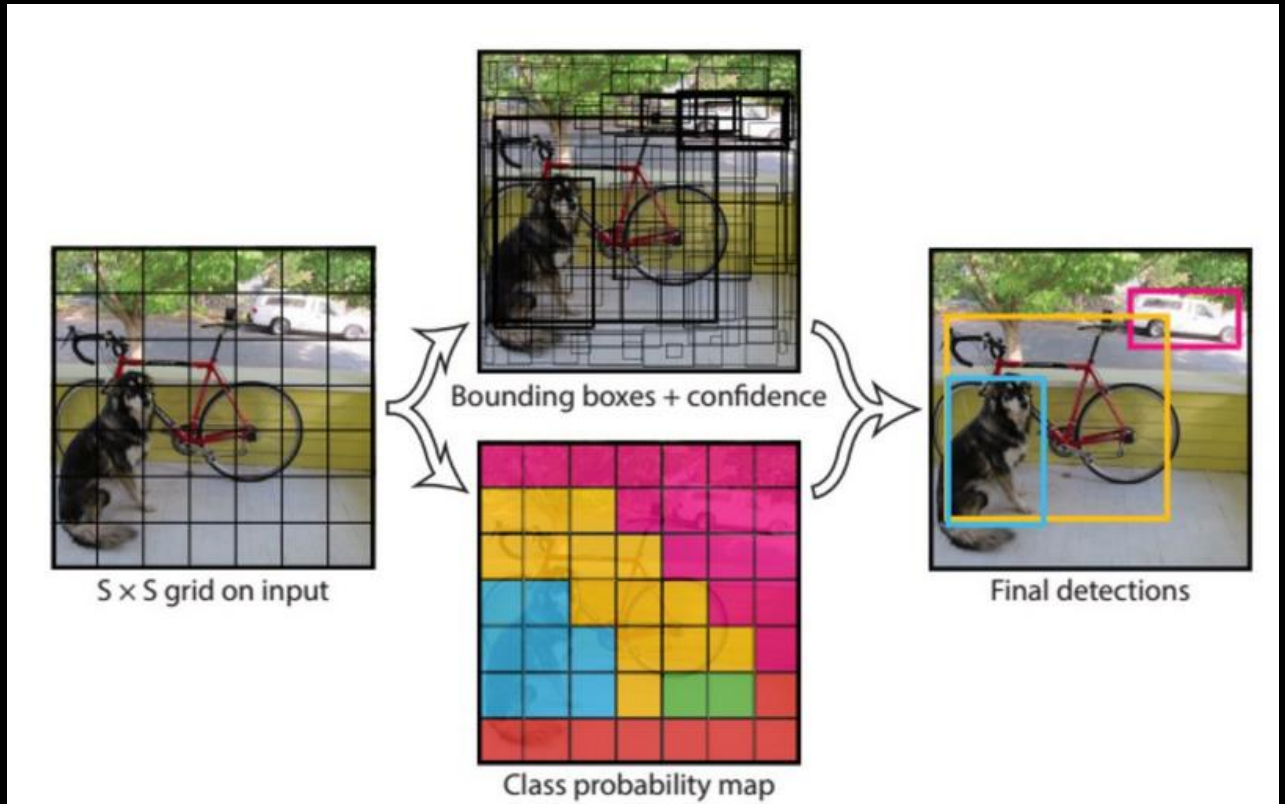

Object is a cat
Classification loss
Refine BB position
Bounding-box regression loss
Object or not object
Classification loss
BB proposal
Bounding-box regression loss
RoI pooling
proposals
Region Proposal Network
feature map
pre-train image-net
CNN
image

OBJECT TRACKING WITH CNN

- Fully-convolutional network tracker (FCNT)



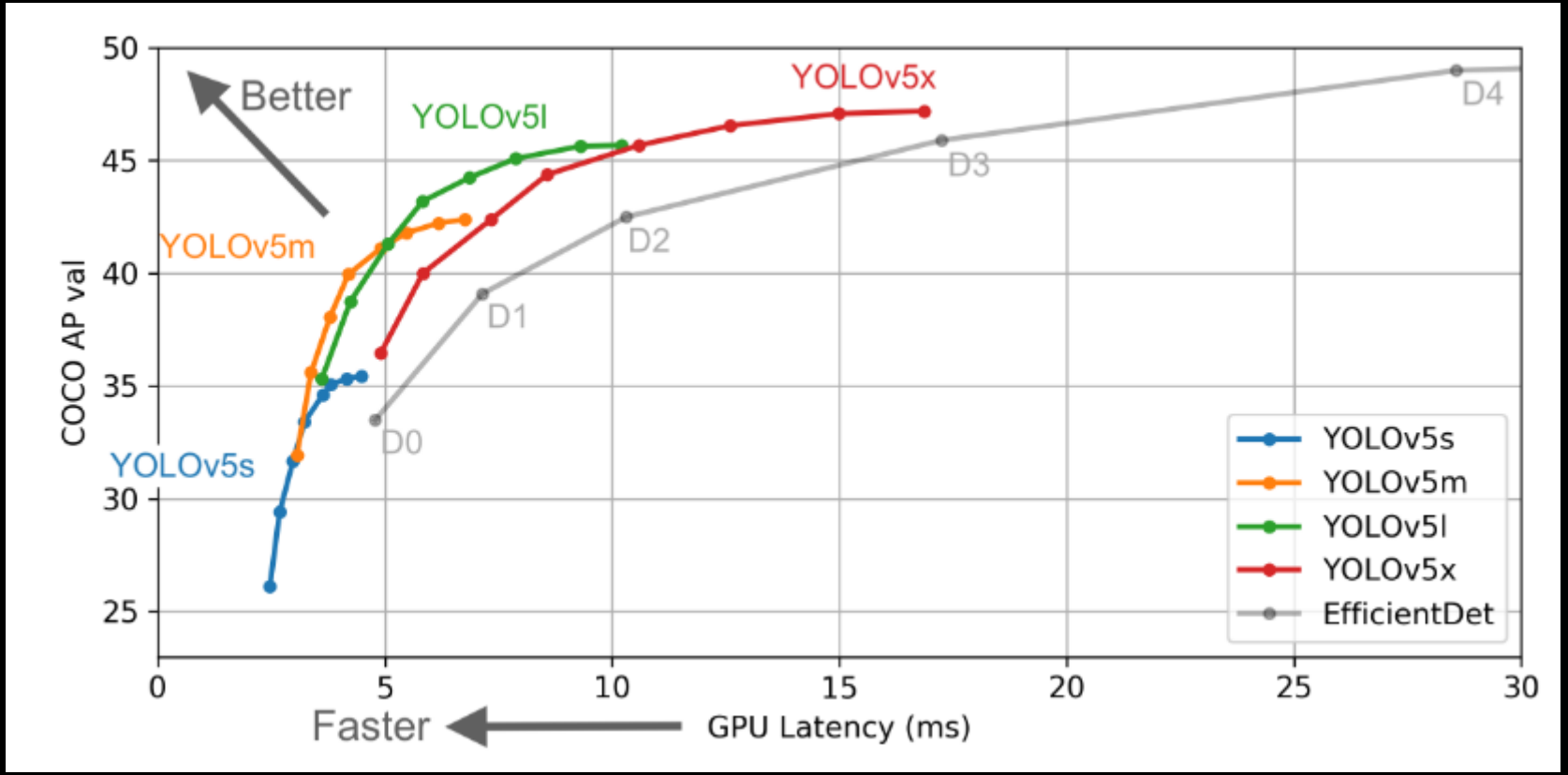
YOU ONLY LOOK ONCE (YOLO)



Source: Redmon et al., You Only Look Once:Unified, Real-Time Object Detection, 2016.

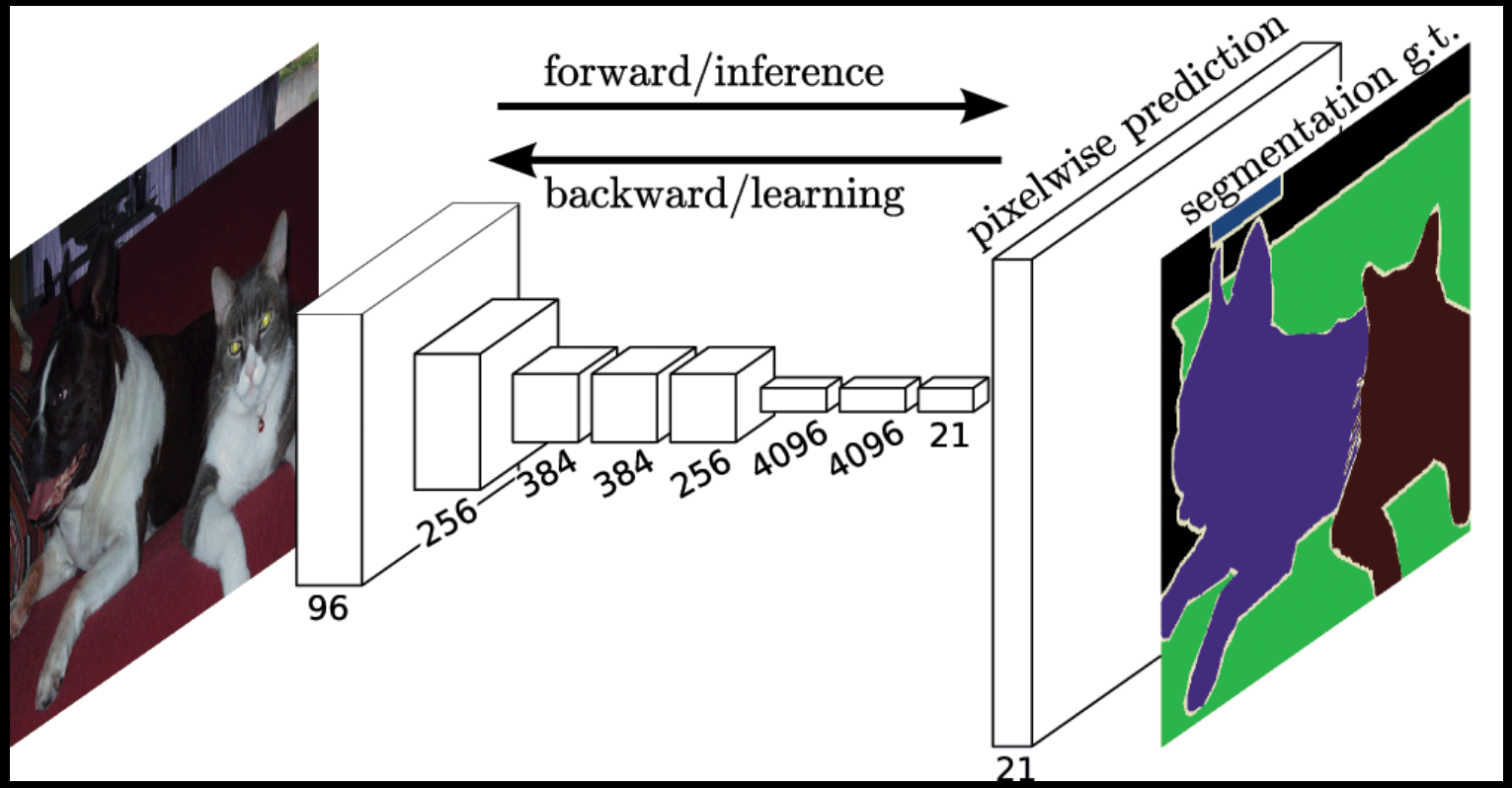
See: https://youtu.be/NM6lrxy0bxs?list=PLrrmP4uhN47Y-hWs7DVfCmLwUACRigYyT
https://pjreddie.com/darknet/yolo/

YOLO V5, JULY 2020



SEMANTIC SEGMENTATION

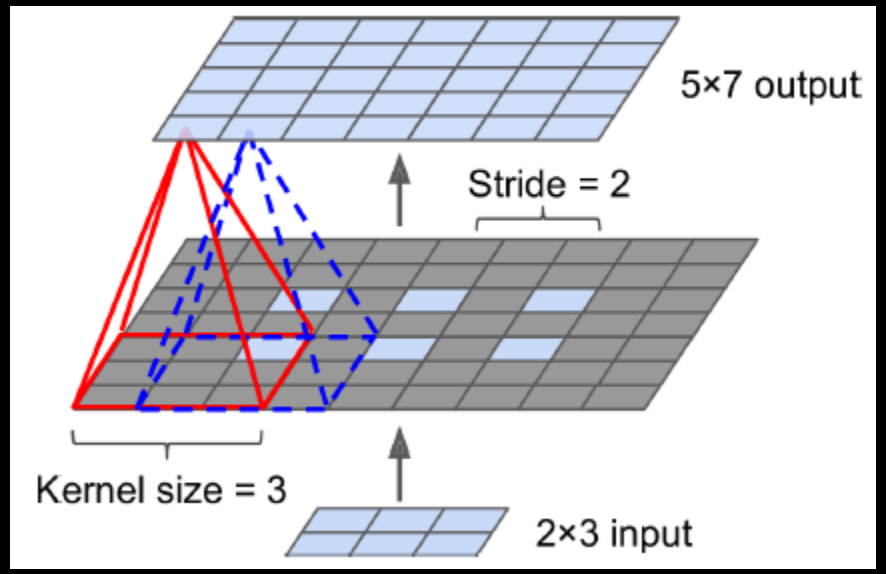
FULLY CONVOLUTIONAL NETWORKS



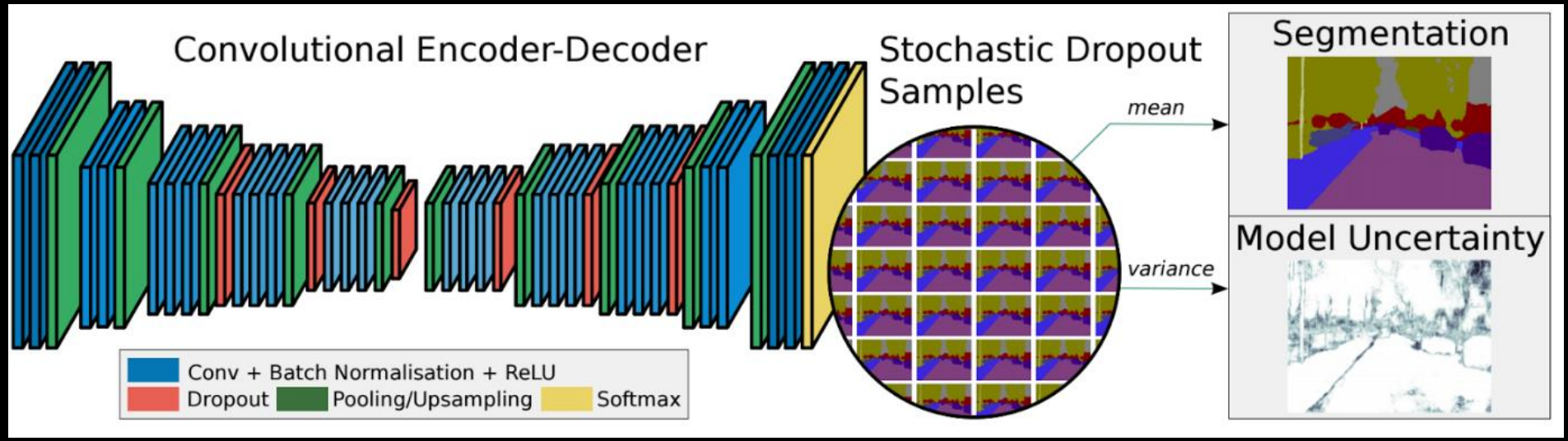
TRANSPOSED CONVOLUTION

- Upsampling layer to recover spatial information

1. stretching
2. filtering



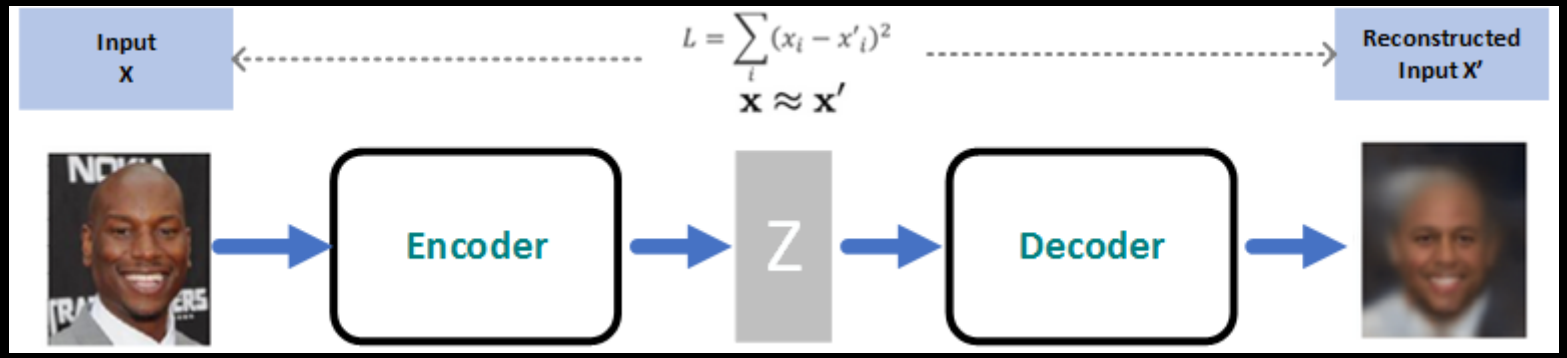
SEGNET


Convolutional Encoder-Decoder
Stochastic Dropout Samples
mean
variance
Segmentation
Model Uncertainty
Conv + Batch Normalisation + ReLU
Dropout
Pooling/Upsampling
Softmax

RECALL DATA AUGMENTATION

AUTOENCODERS

- Generating augmented data



CONDITIONAL VARIATIONAL AUTOENCODER

- Output determined by latent variables, chance and metadata.