Fast RCNN

Fast R-CNN

Ross Girshick
Microsoft Research
rbg@microsoft.com

Abstract

This paper proposes a Fast Region-based Convolutional Network method (Fast R-CNN) *for object detection. Fast R-CNN builds on previous work to efficiently classify object proposals using deep convolutional networks. Compared to previous work, Fast R-CNN employs several innovations to improve training and testing speed while also increasing detection accuracy. Fast R-CNN trains the very deep VGG16 network 9× faster than R-CNN, is 213× faster at test-time, and achieves a higher mAP on PASCAL VOC 2012. Compared to SPPnet, Fast R-CNN trains VGG16 3× faster, tests 10× faster, and is more accurate. Fast R-CNN is implemented in Python and C++ (using Caffe) and is available under the open-source MIT License at* https://github.com/rbgirshick/fast-rcnn.

while achieving top accuracy on PASCAL VOC 2012 [7] with a mAP of 66% (vs. 62% for R-CNN).[1]

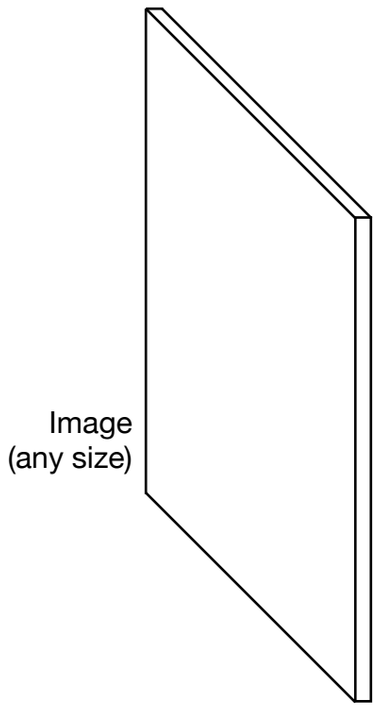
1.1. R-CNN and SPPnet

The Region-based Convolutional Network method (R-CNN) [9] achieves excellent object detection accuracy by using a deep ConvNet to classify object proposals. R-CNN, however, has notable drawbacks:

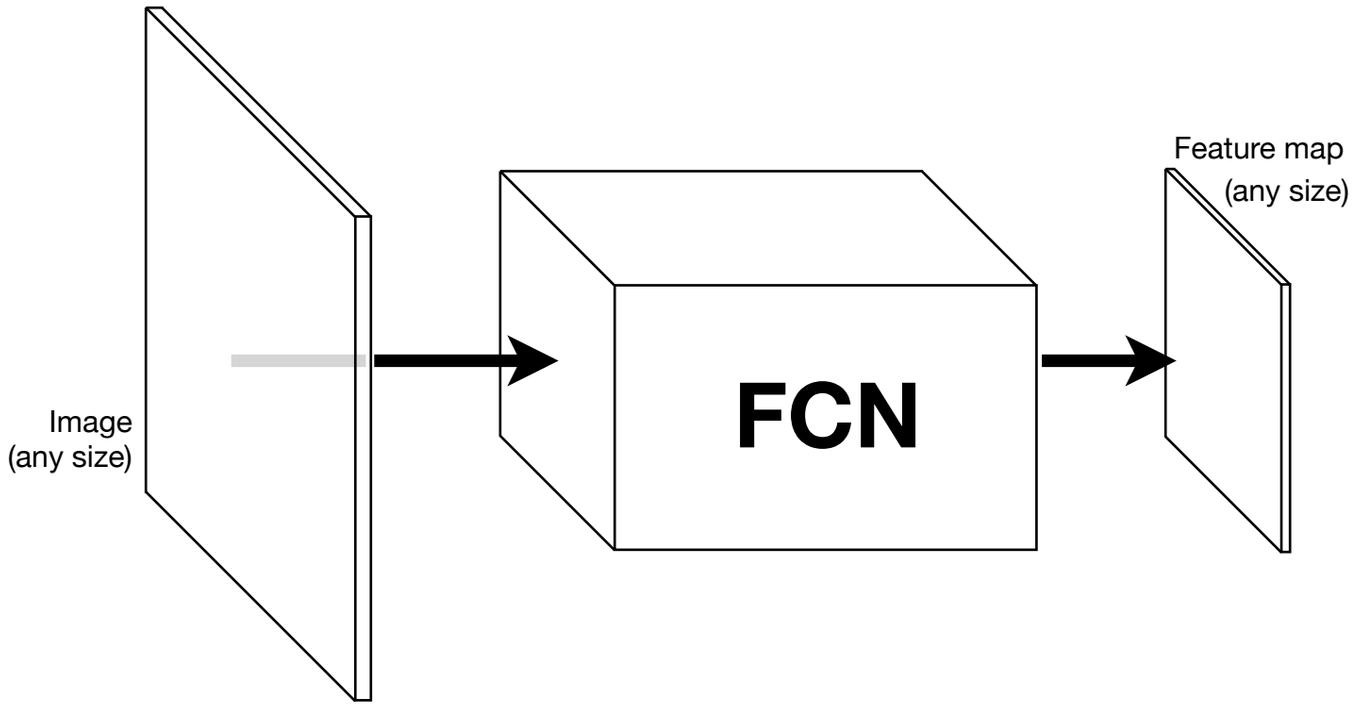
1. **Training is a multi-stage pipeline.** R-CNN first fine-tunes a ConvNet on object proposals using log loss. Then, it fits SVMs to ConvNet features. These SVMs act as object detectors, replacing the softmax classifier learnt by fine-tuning. In the third training stage, bounding-box regressors are learned.
2. **Training is expensive in space and time.** For SVM

2015

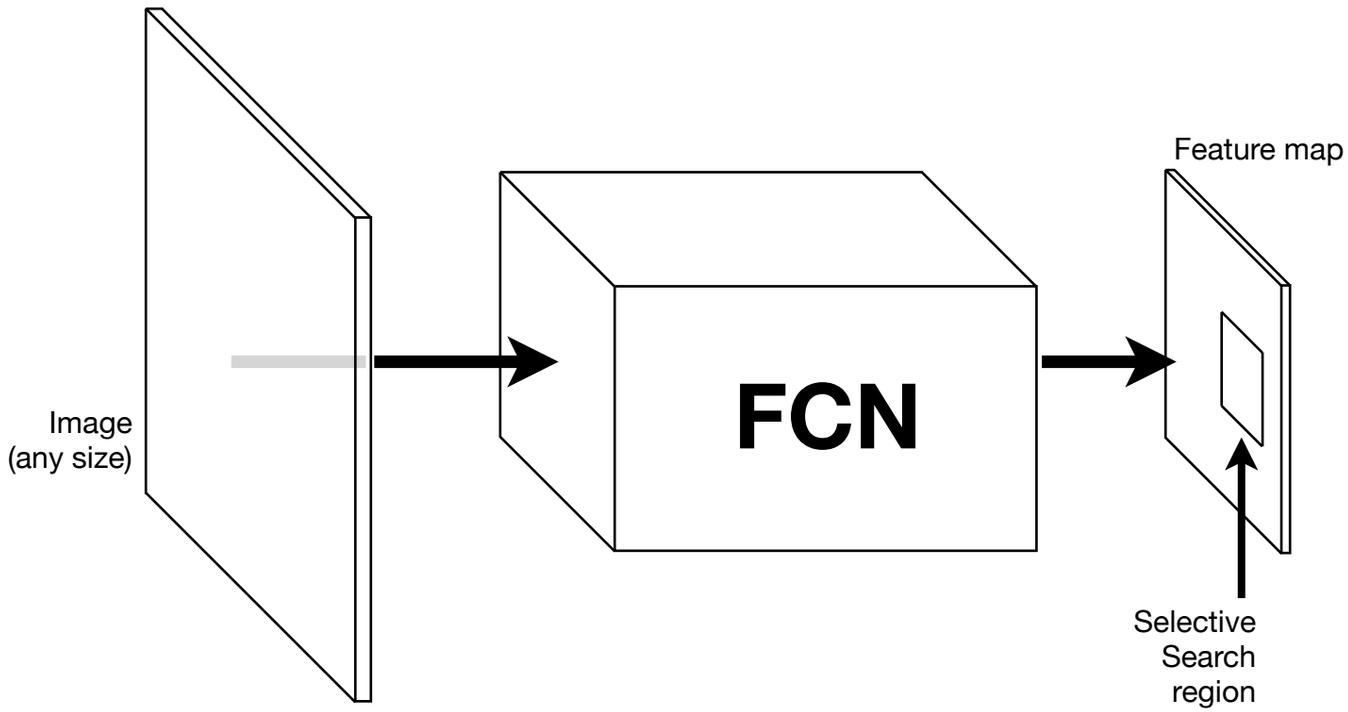
Fast RCNN



Fast RCNN



Fast RCNN



Fast RCNN

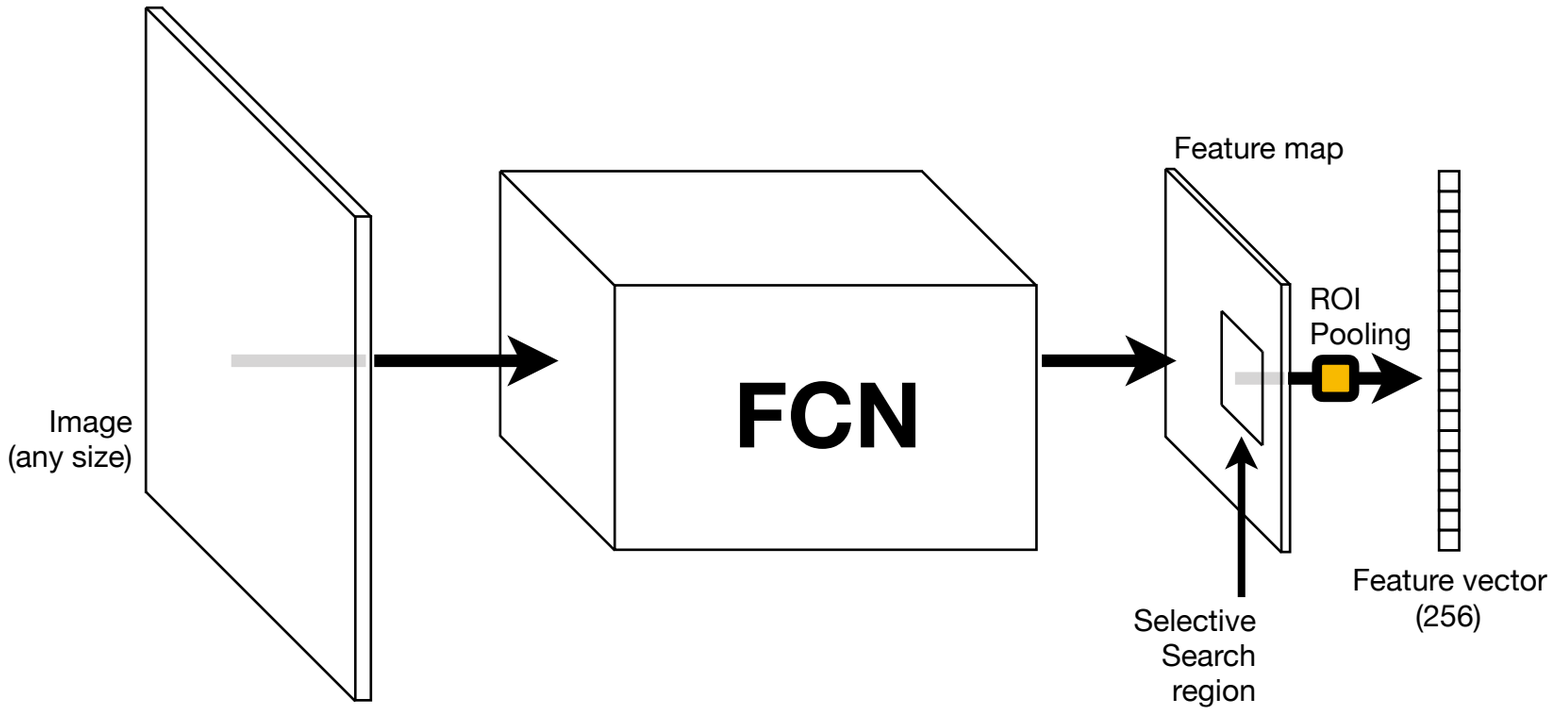
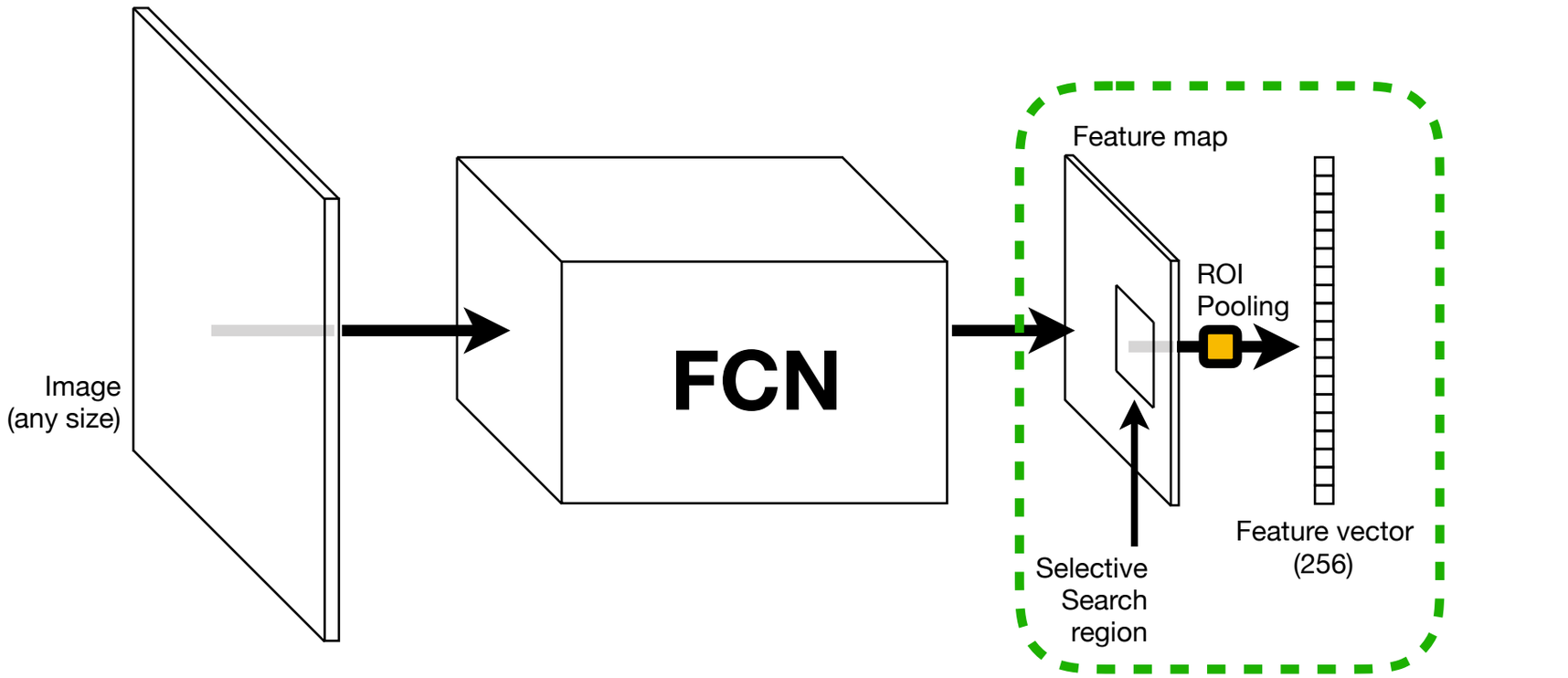


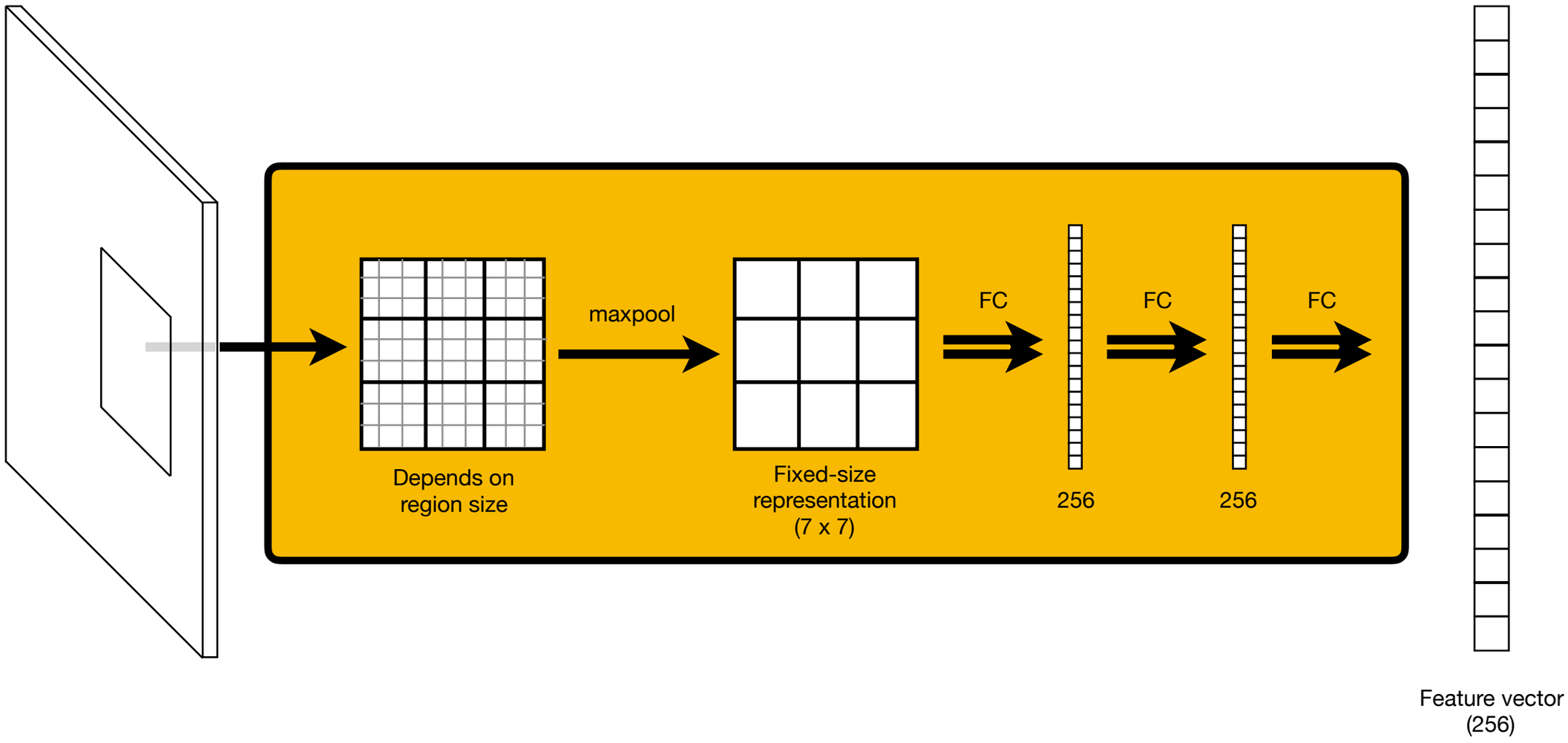
Image
(any size)
FCN
Feature map
ROI
Pooling
Feature vector
(256)
Selective
Search
region

Fast RCNN



ROI Pooling Layer

ROI Pooling Layer



Fast RCNN

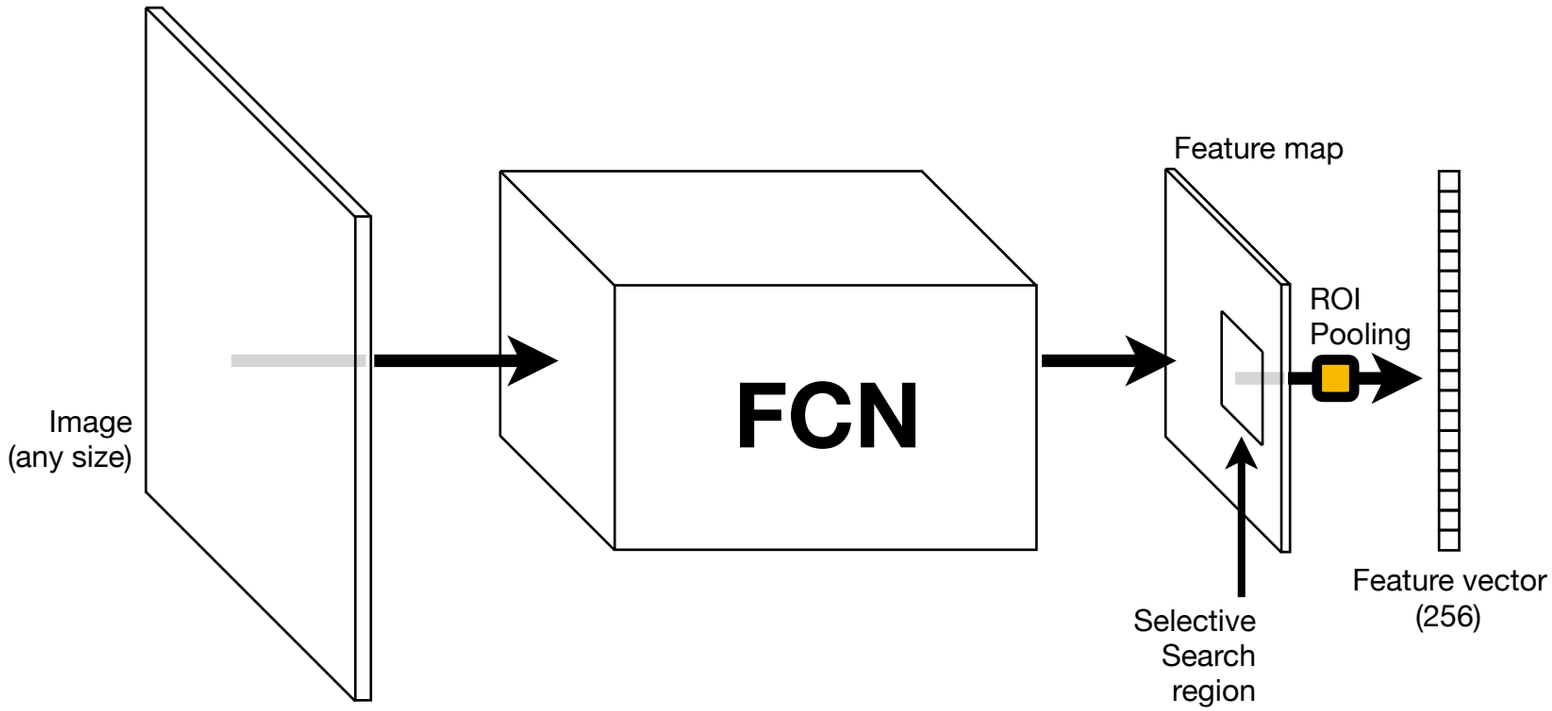


Image
(any size)
FCN
Feature map
ROI
Pooling
Feature vector
(256)
Selective
Search
region

Fast RCNN

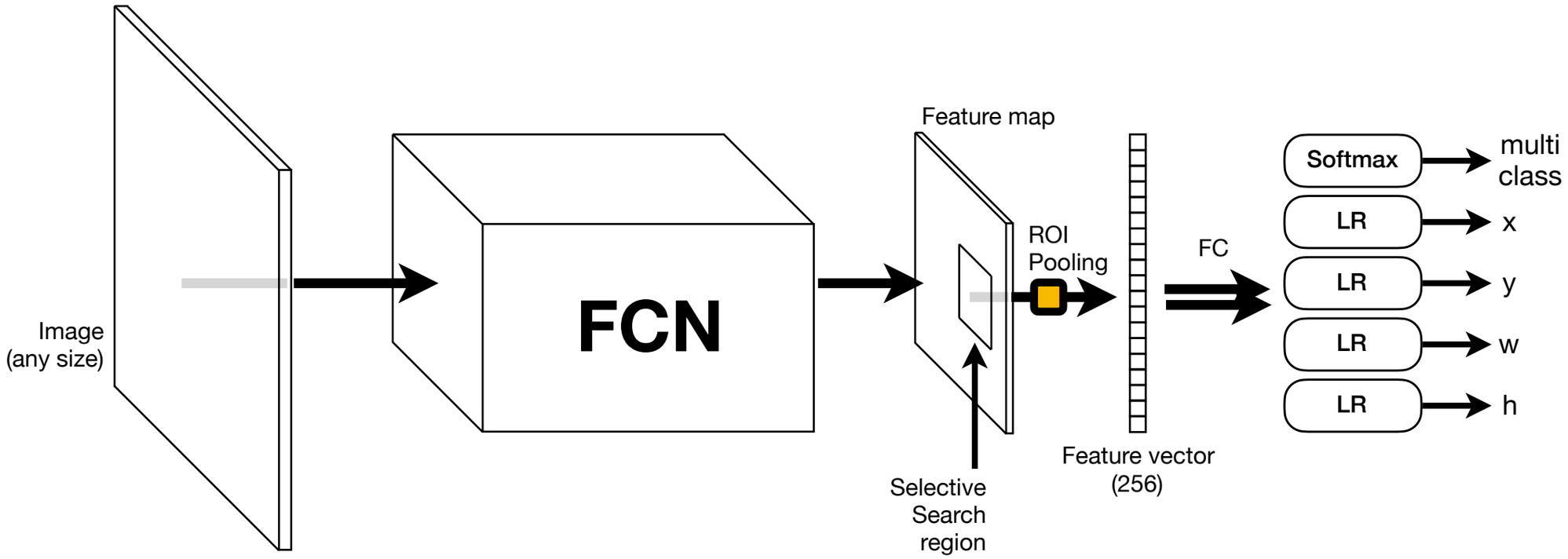
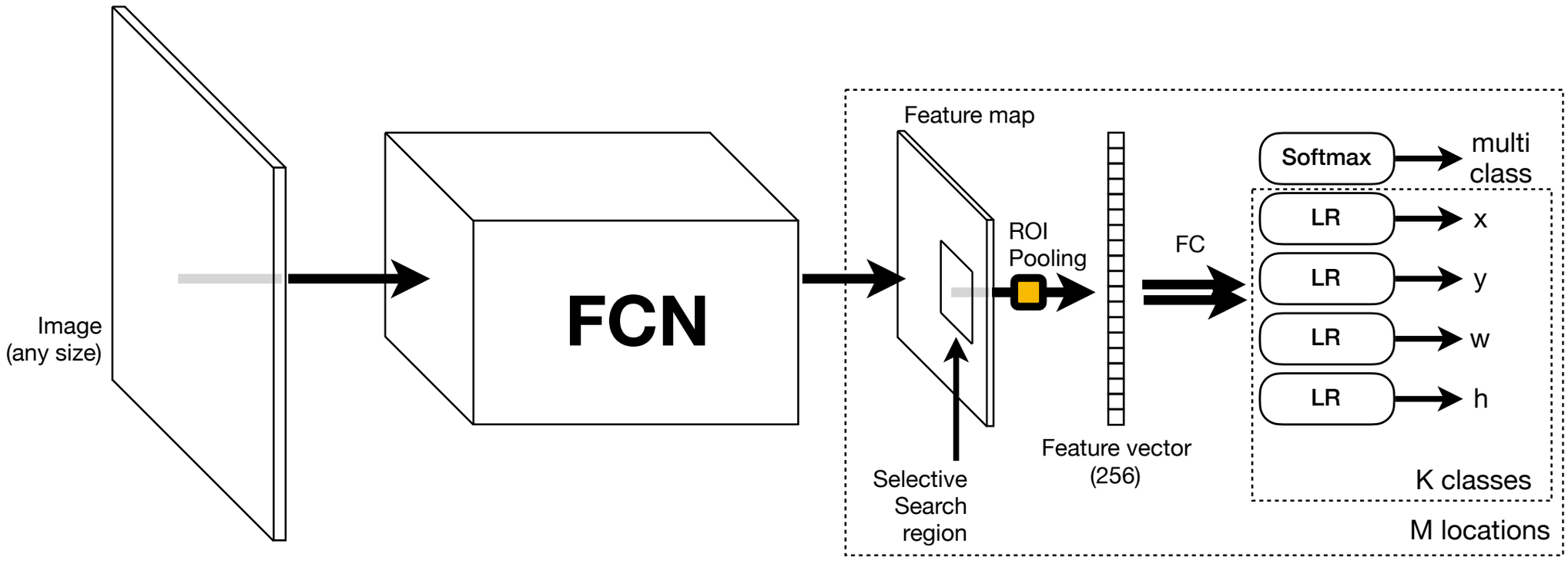


Image
(any size)
FCN
Feature map
ROI
Pooling
Selective
Search
region
Feature vector
(256)
FC
Softmax
multi
class
LR
x
LR
y
LR
w
LR
h

Fast RCNN



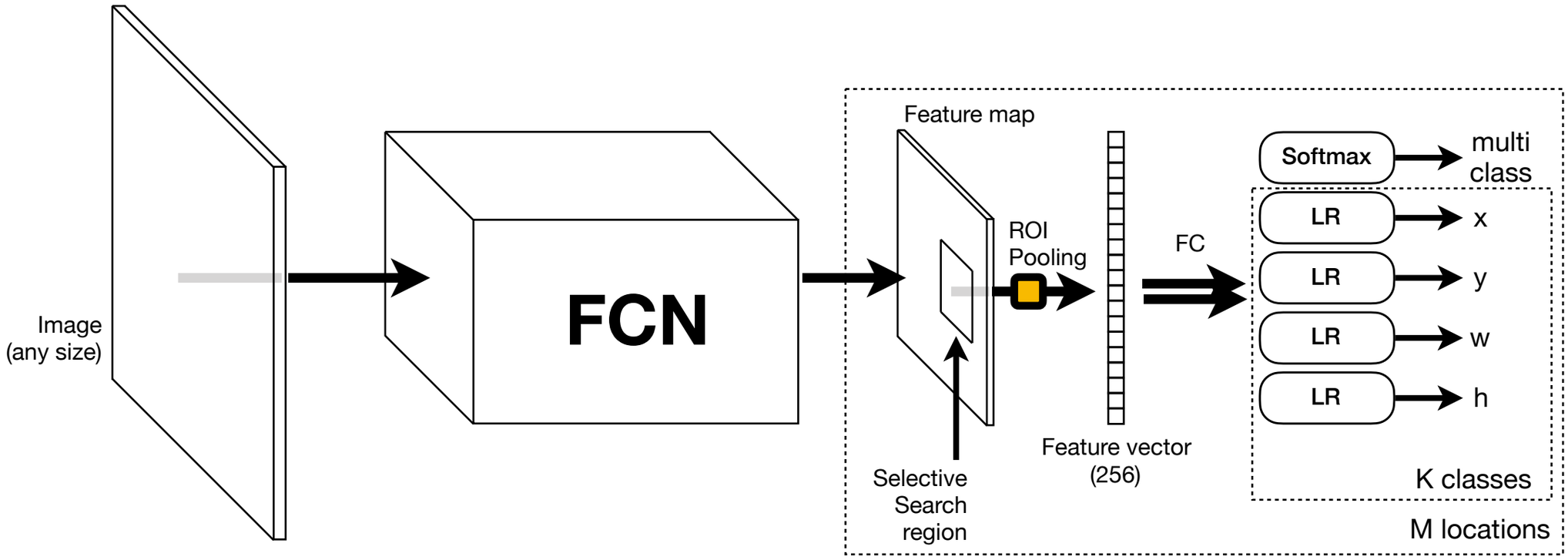
Complete.

Important Concepts

- FCN conversion for variable sized input

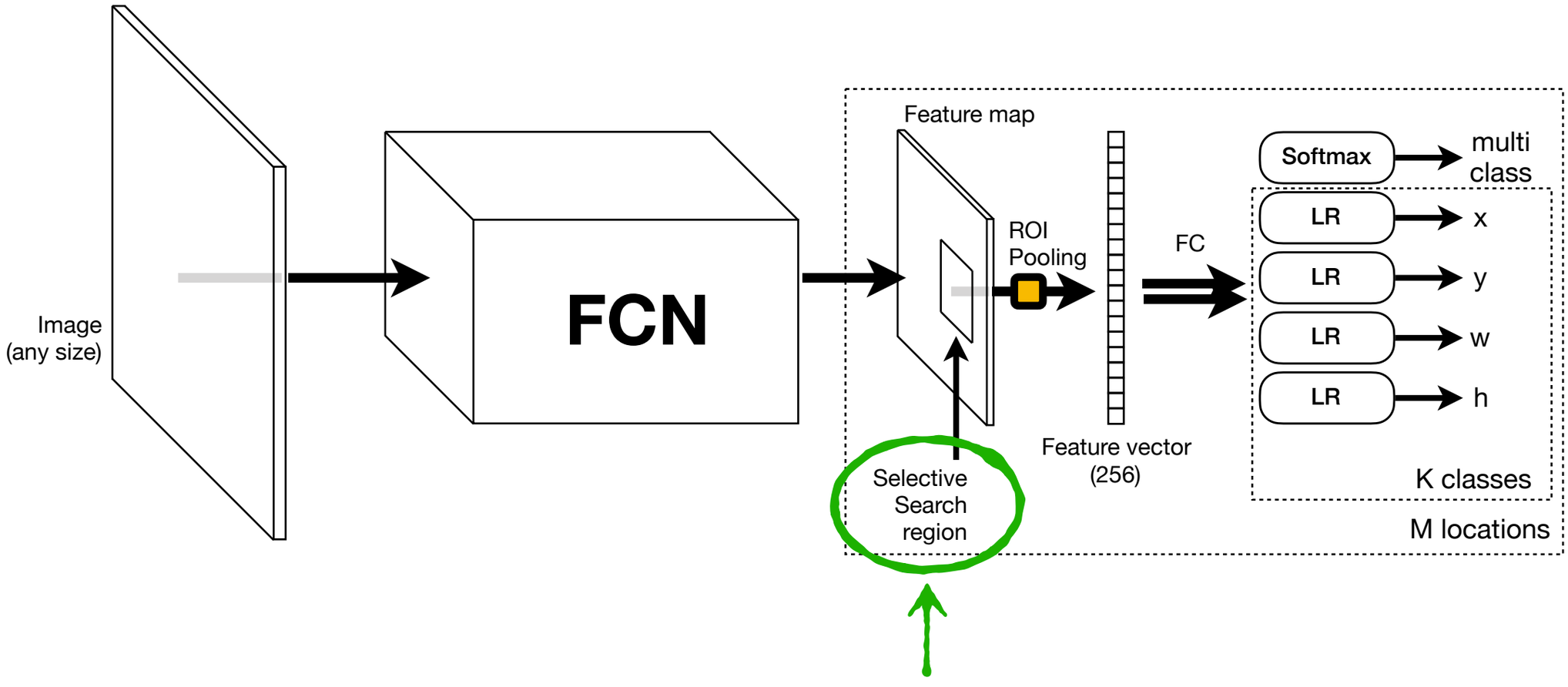
- ROI Pooling

Fast RCNN



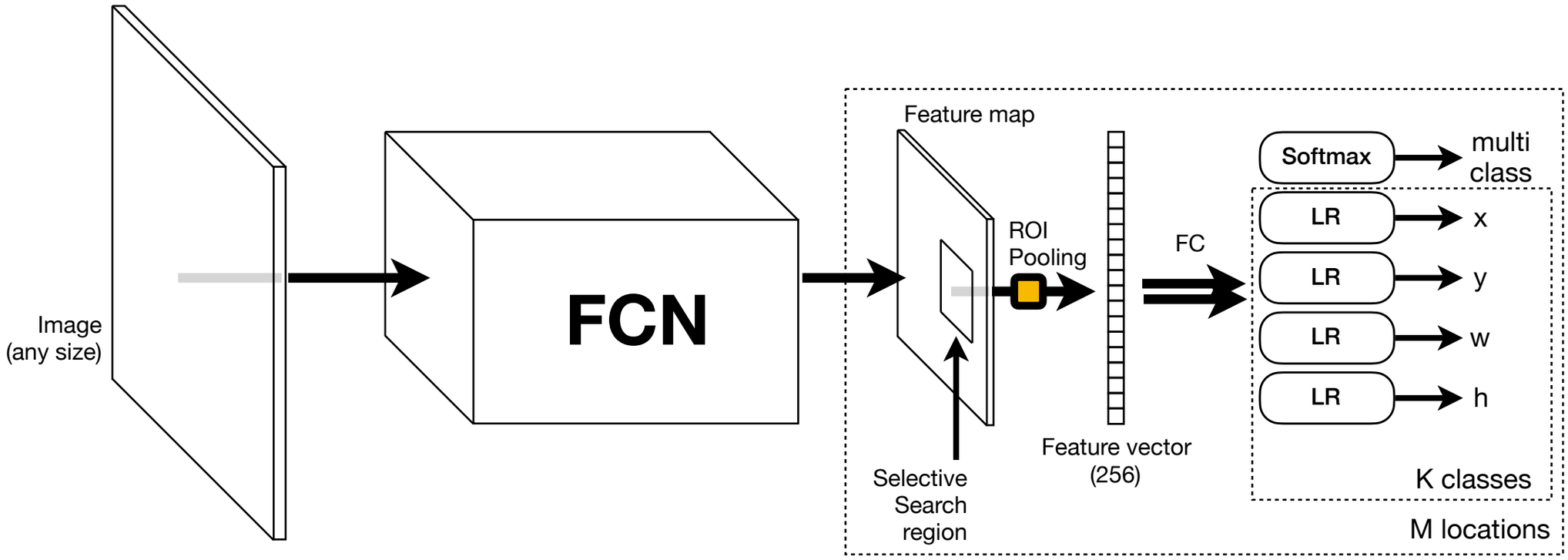
It turns out that Fast RCNN is not really that fast.
Why?

Fast RCNN



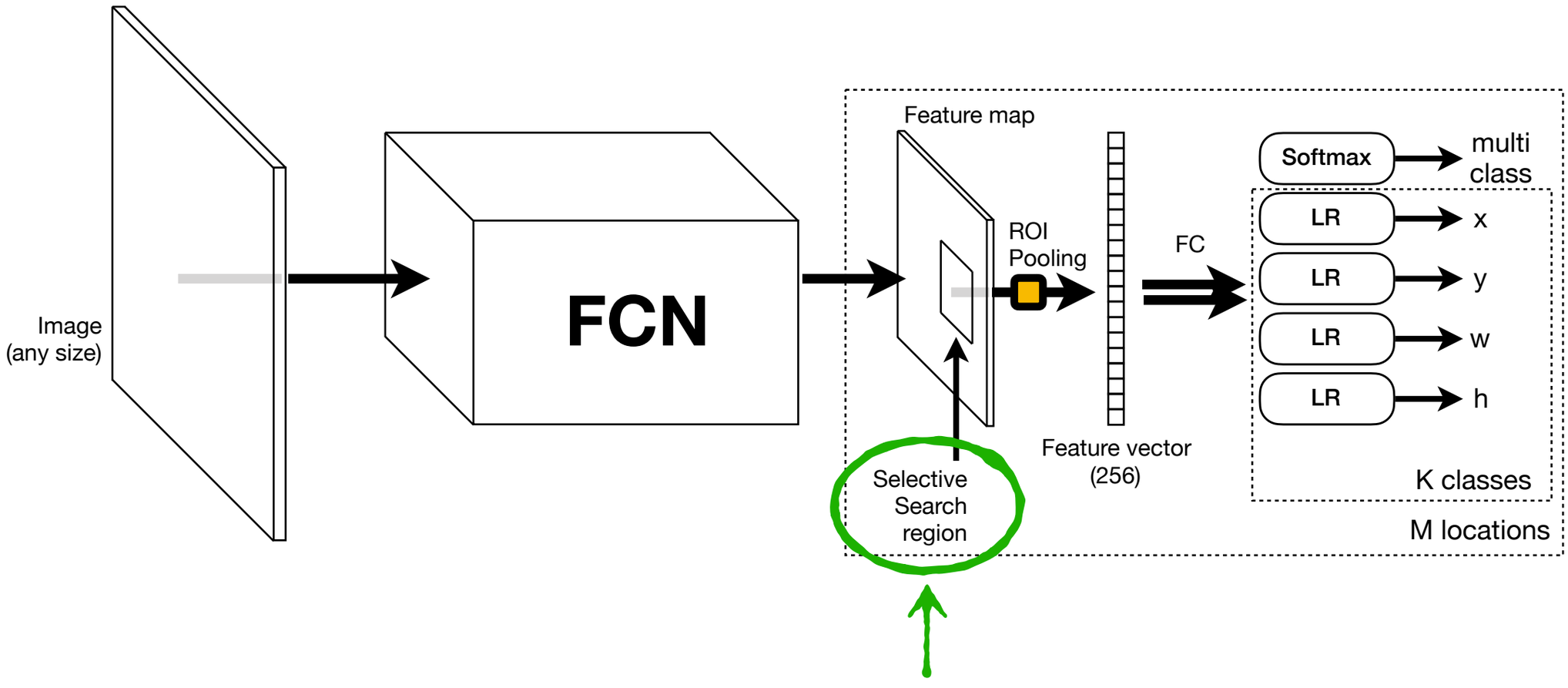
Selective search is really slow!

Fast RCNN



How can we make Fast-RCNN FASTER???

Fast RCNN



Use a deep network to generate proposals!

Fast RCNN

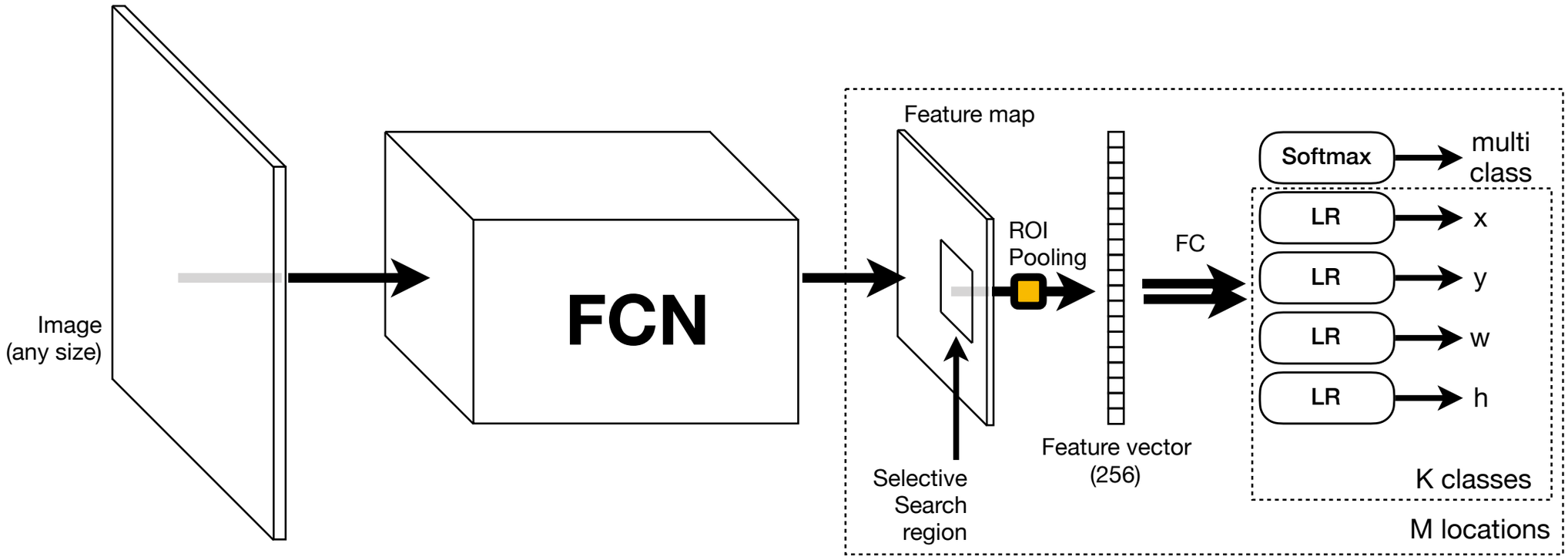
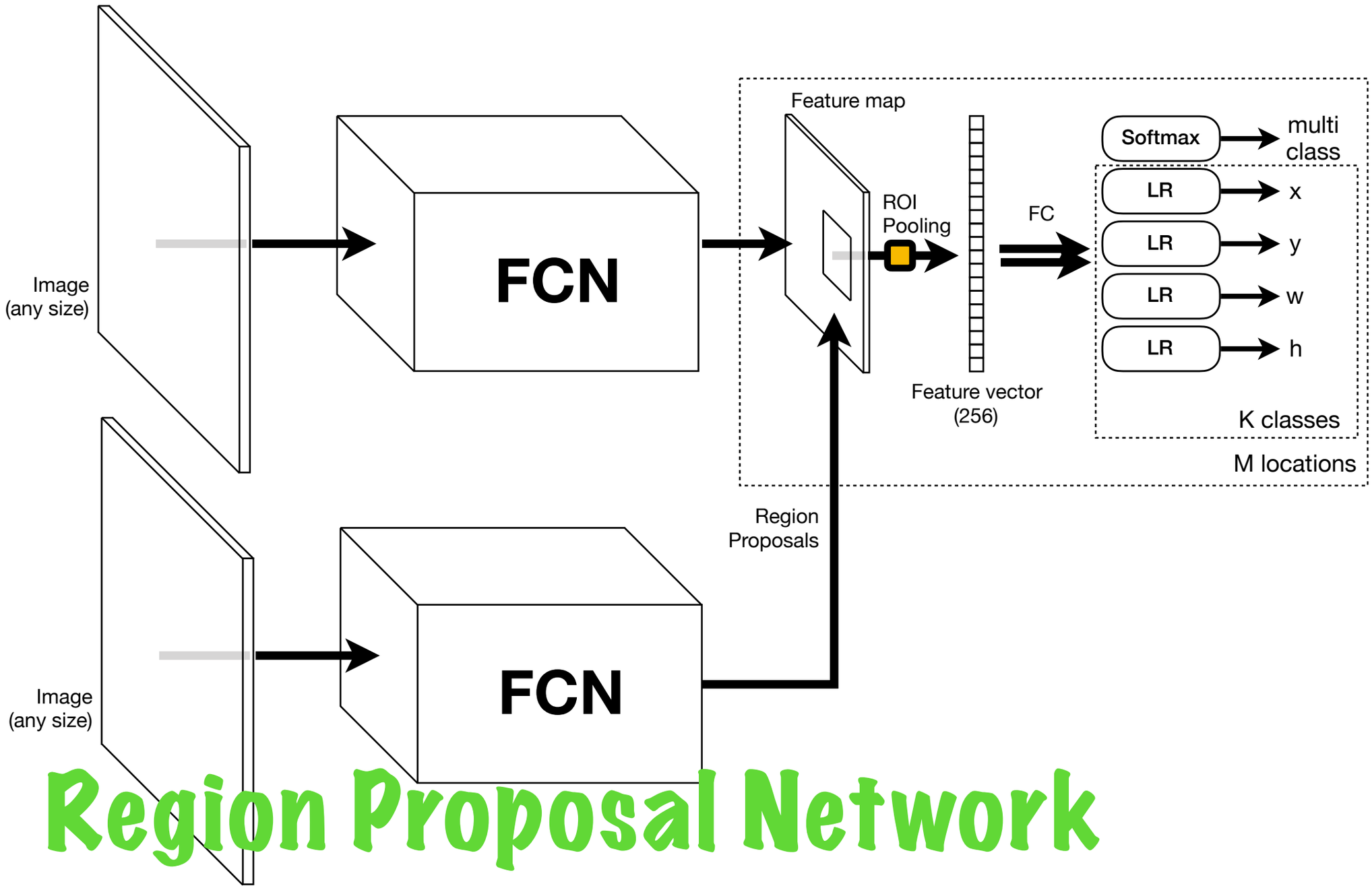


Image
(any size)
FCN
Feature map
ROI
Pooling
FC
Feature vector
(256)
Softmax
multi
class
LR
x
LR
y
LR
w
LR
h
K classes
Selective
Search
region
M locations

Faster

Fast RCNN



Region Proposal Network