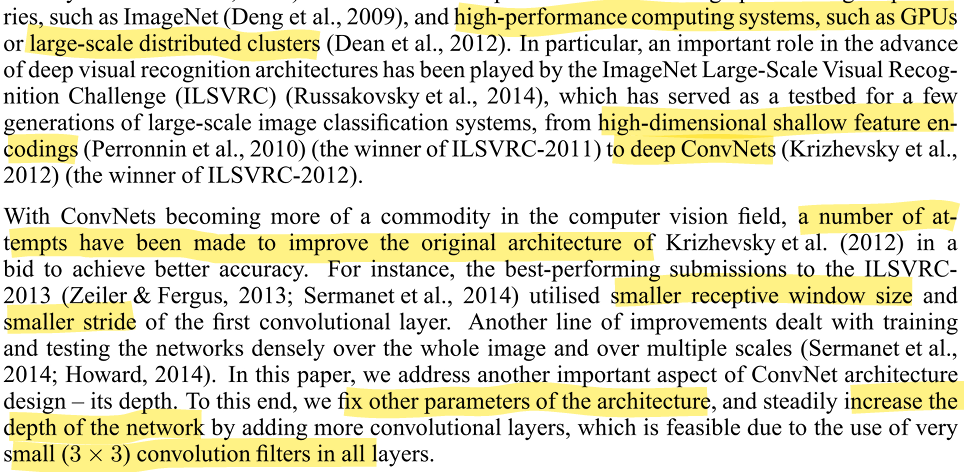
# VGG Paper

Visual Geometry Group

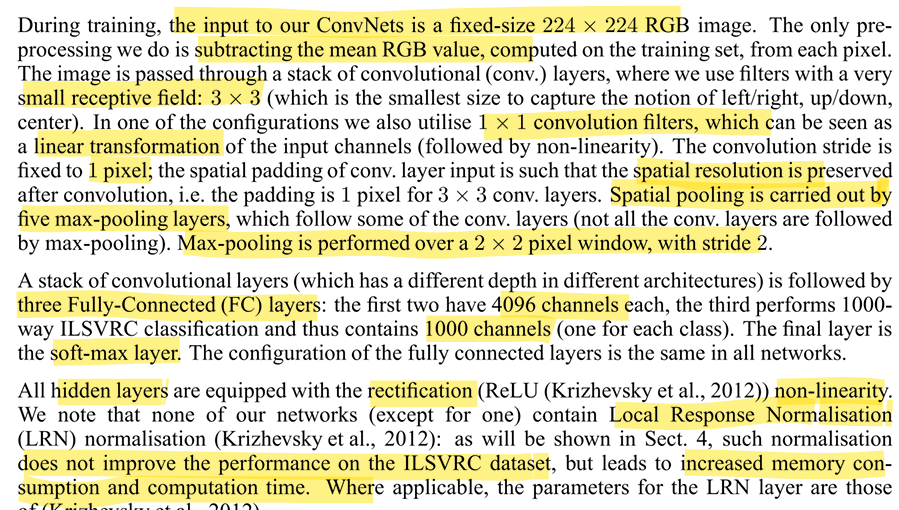
Main Contribution :

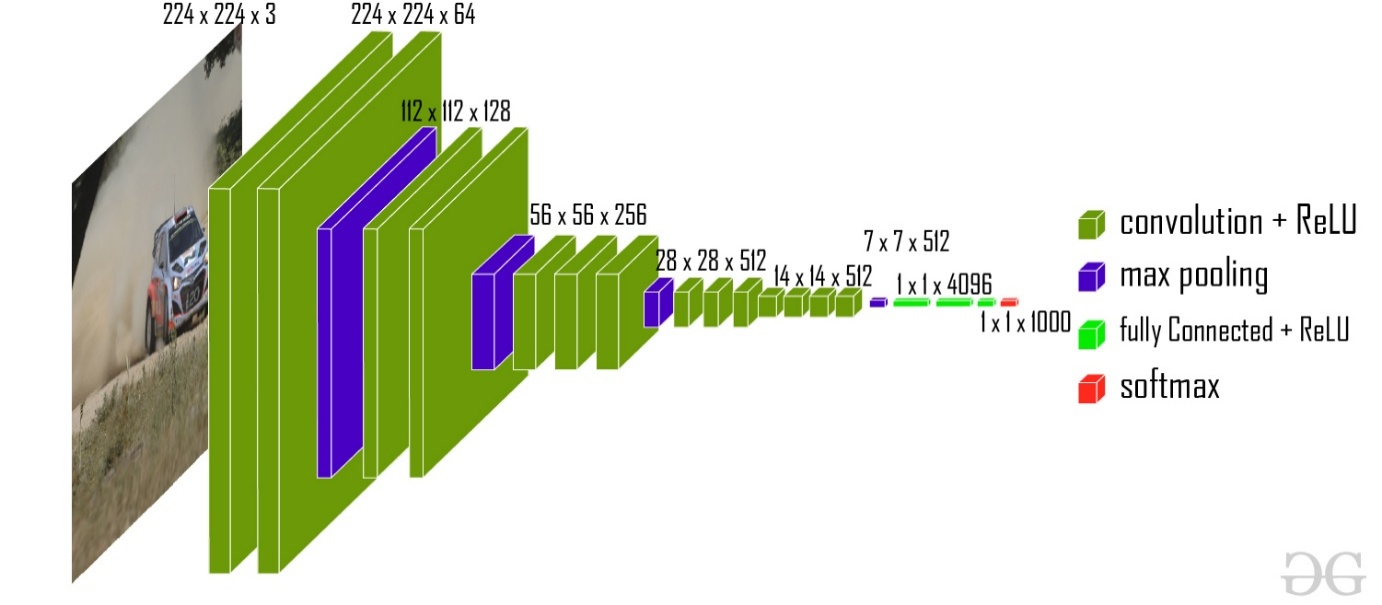
* (3x3) convolution filters
* 16-19 weight layers

What do we see in this paper ?

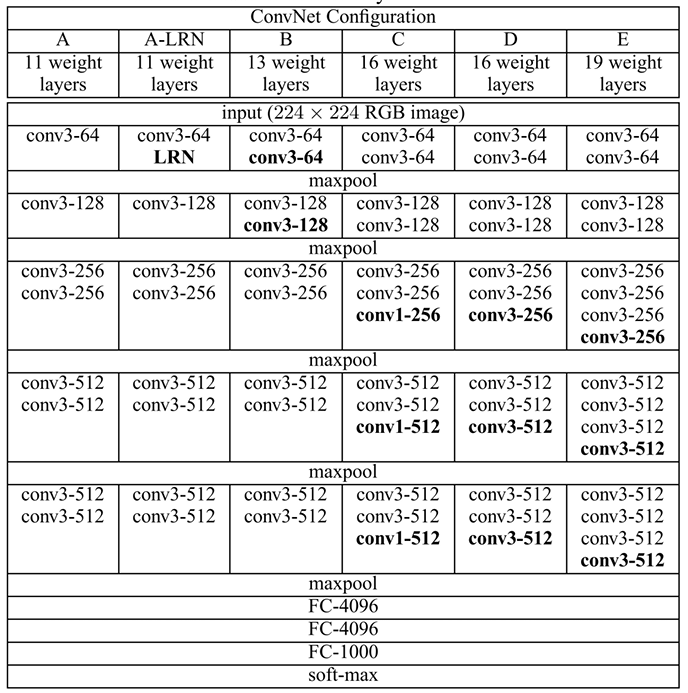


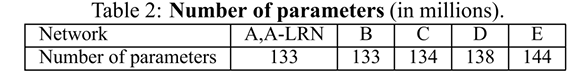
Architecture :

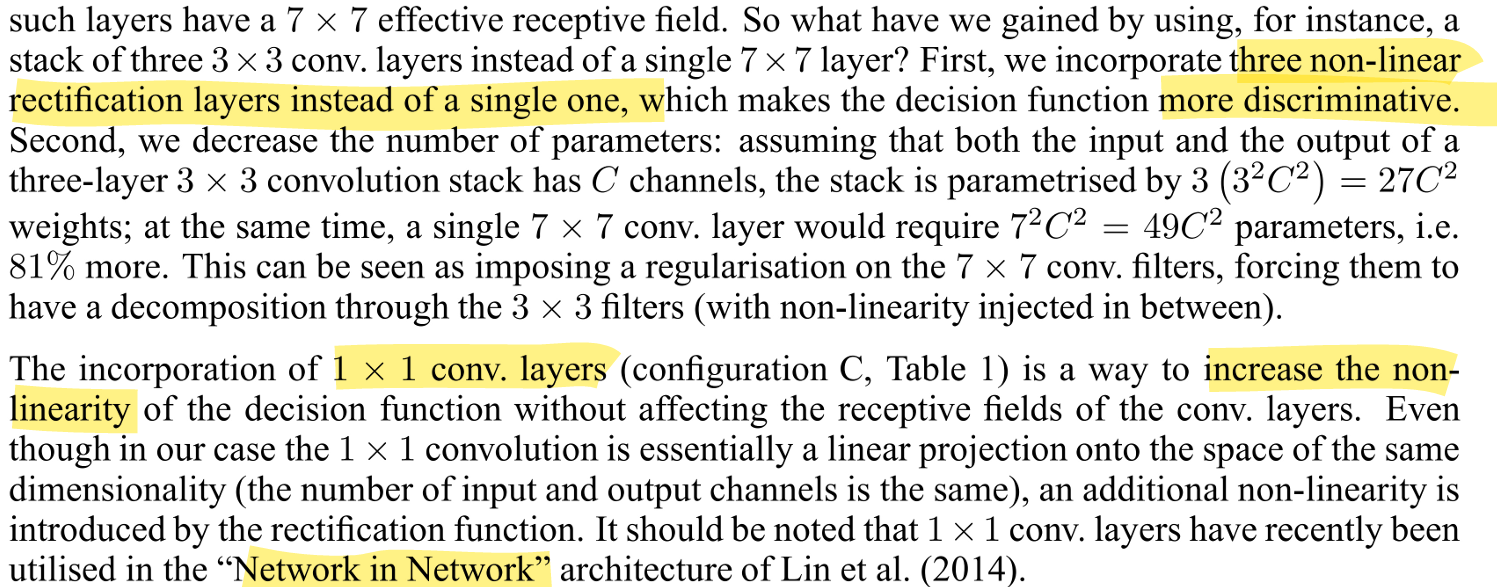


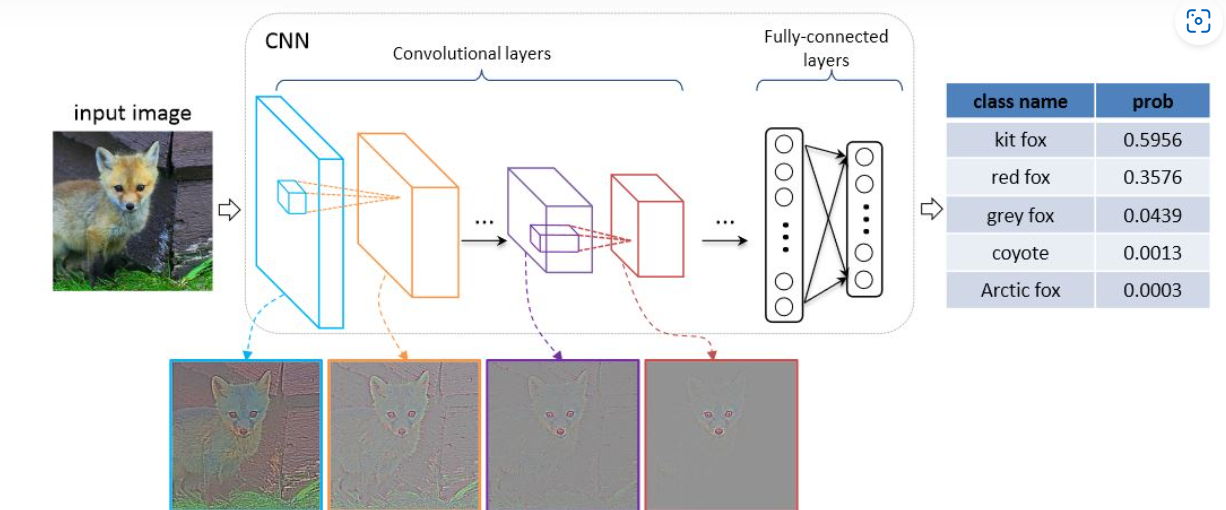


16 weight layers model

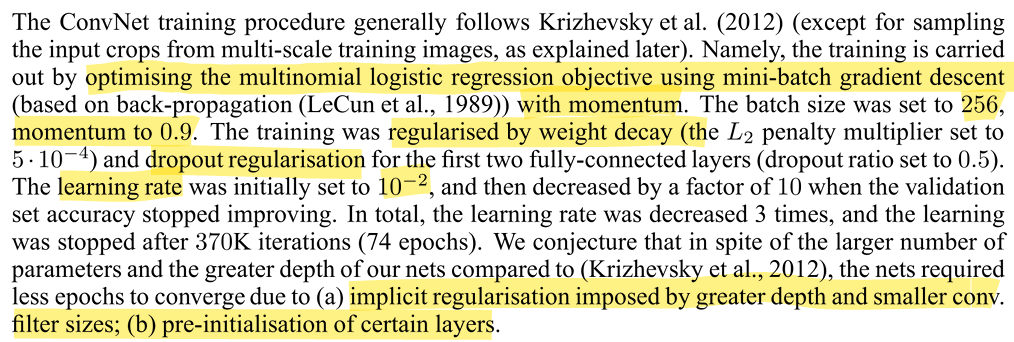


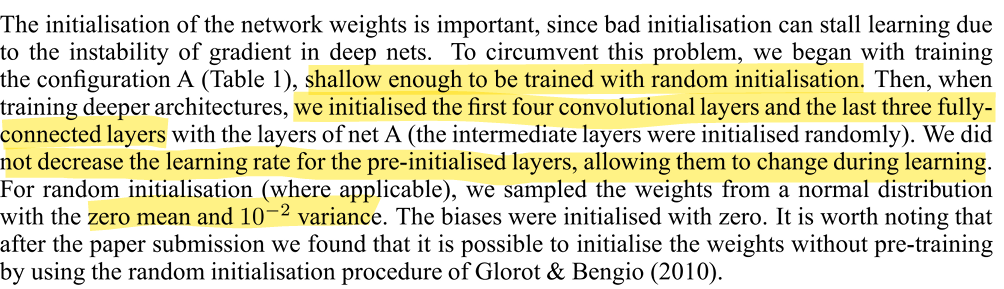






How did the data get trained ?

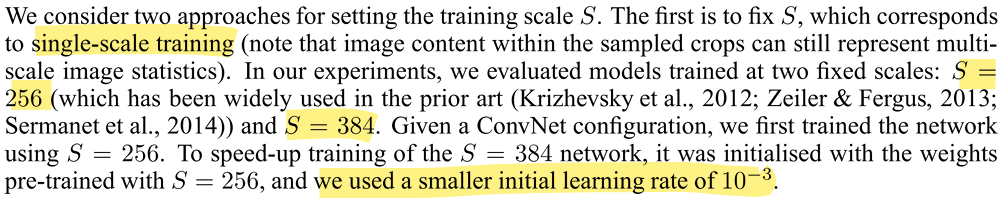




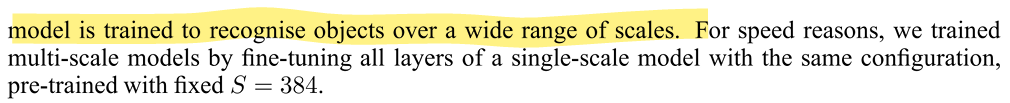
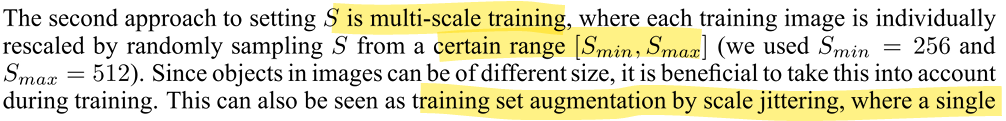
Training Image Size :  
Let S be the smallest side of an isotropically-rescaled training image

There are two approaches for training the images

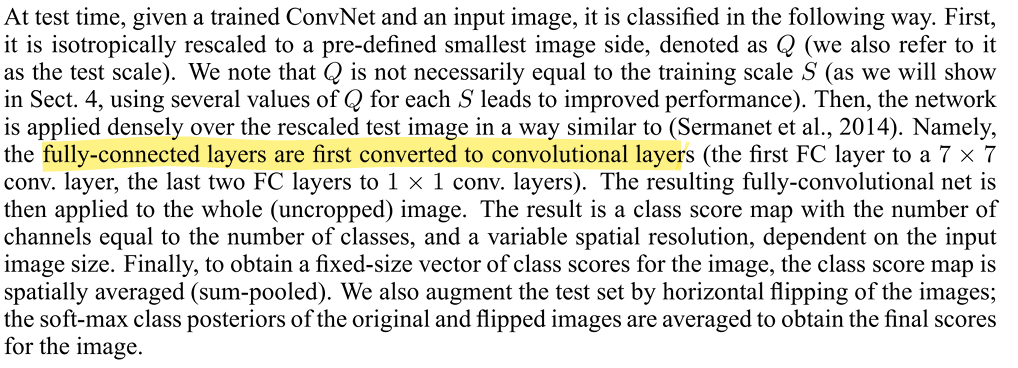
1.

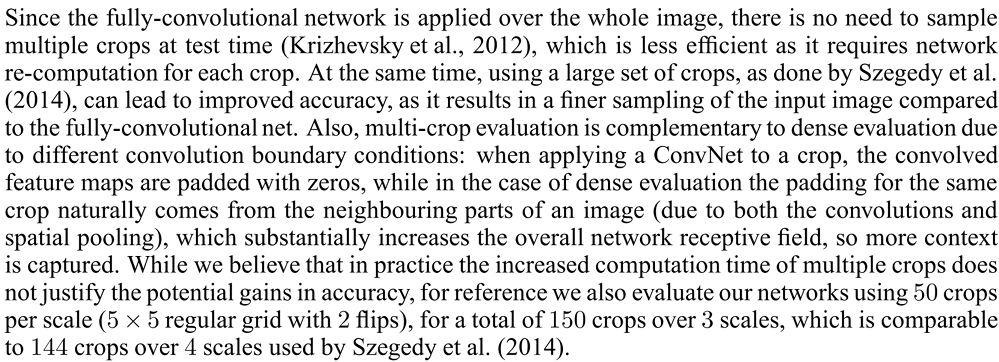


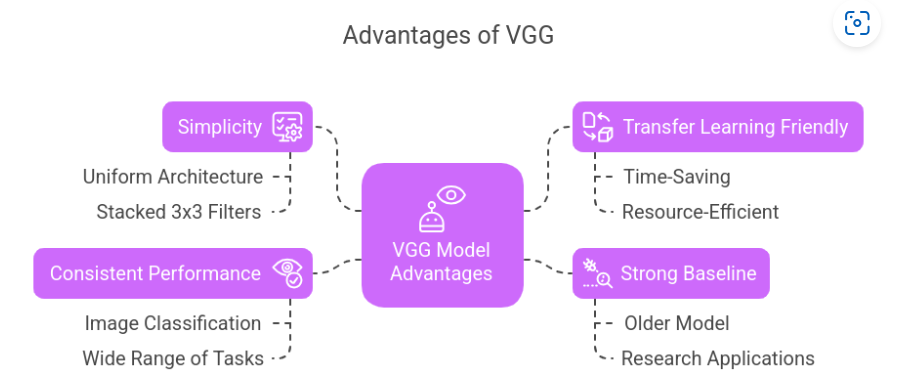
2.



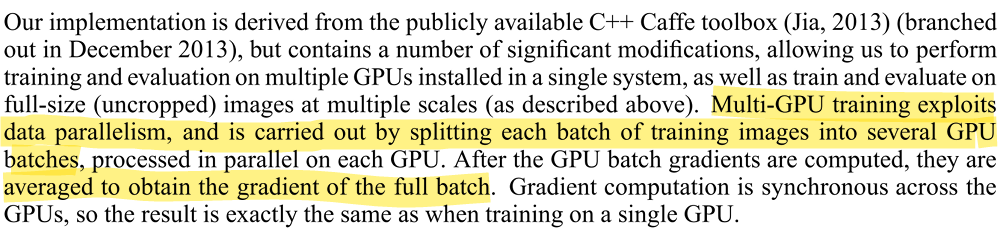
Training :



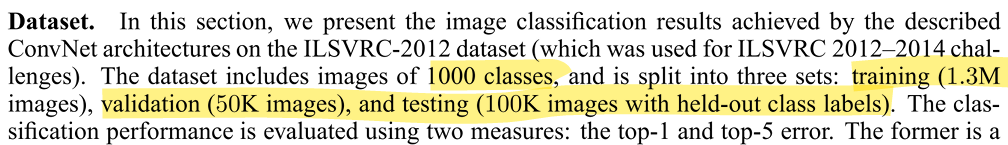




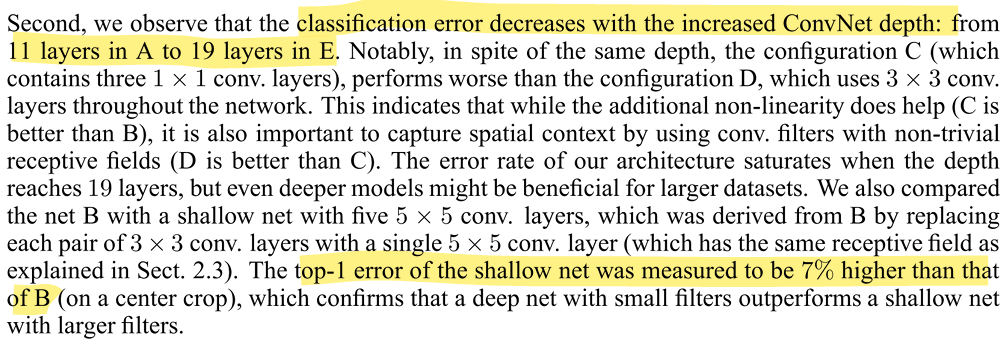
Implementation :

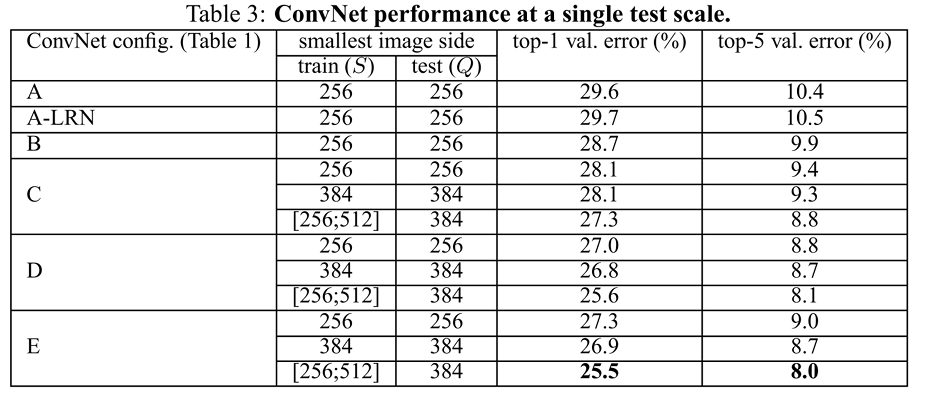


Classification Experiments :

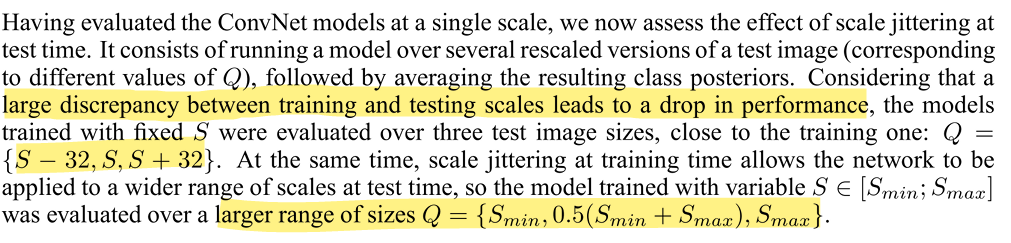


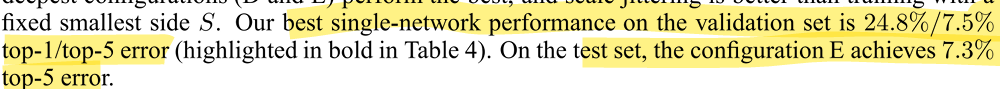
Single Scale Evaluation :

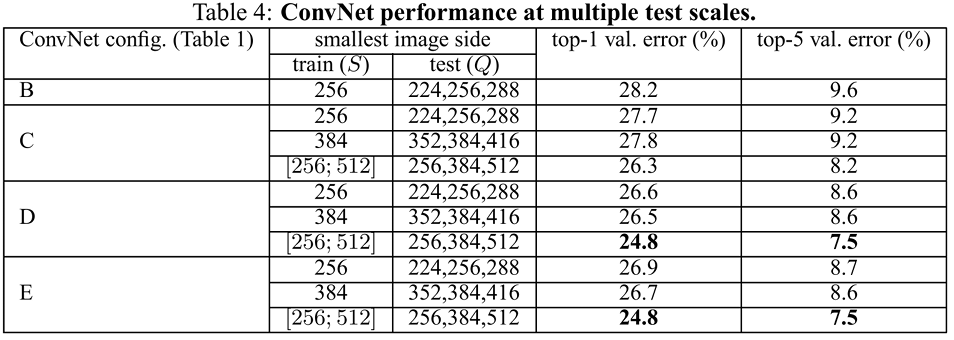




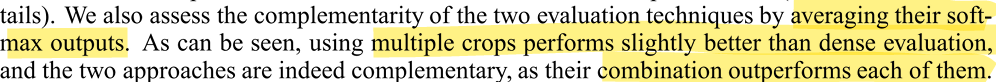
Multi-Scale Evaluation :

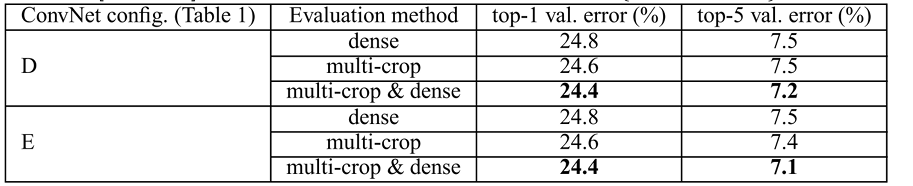




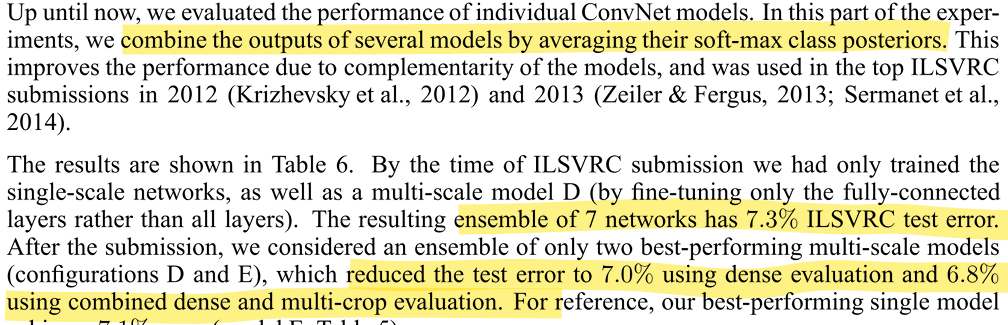


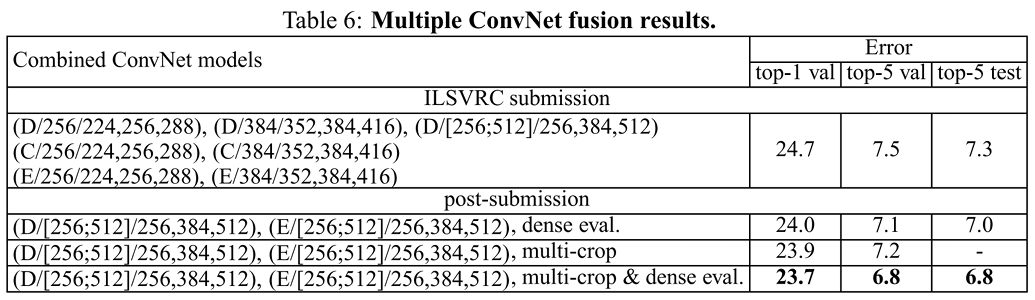
Multi-Crop Evaluation :

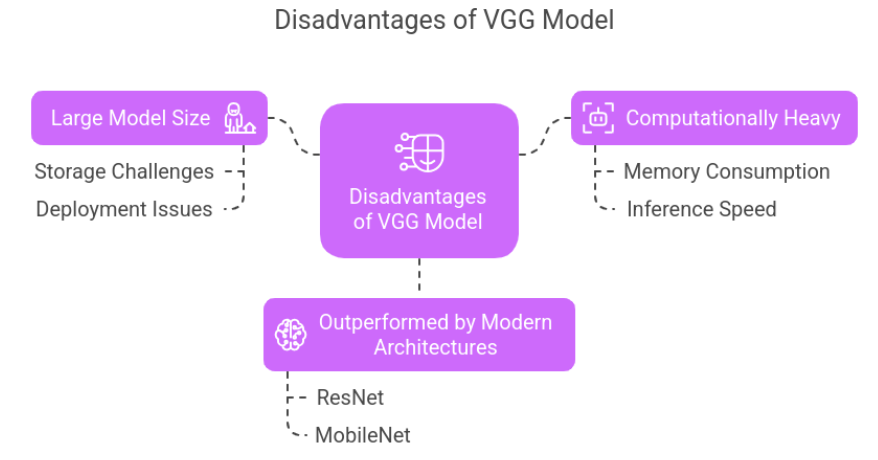




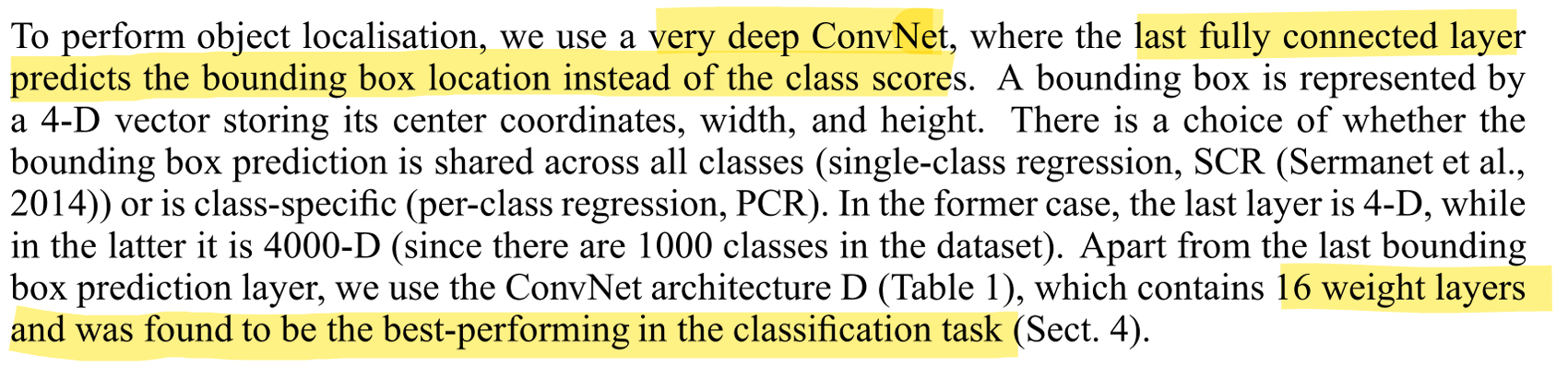
ConvNet fusion :



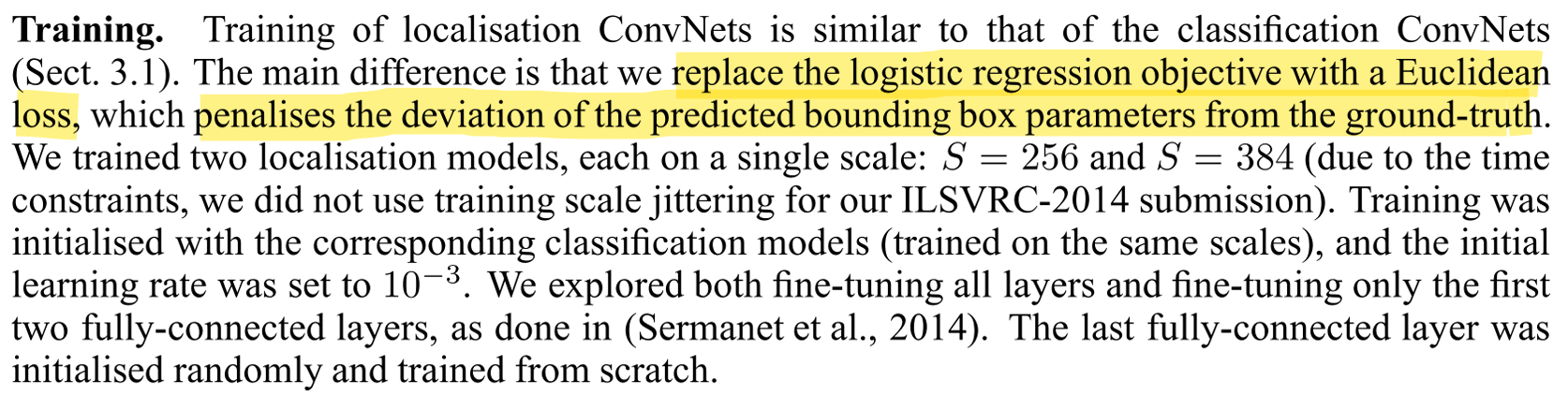




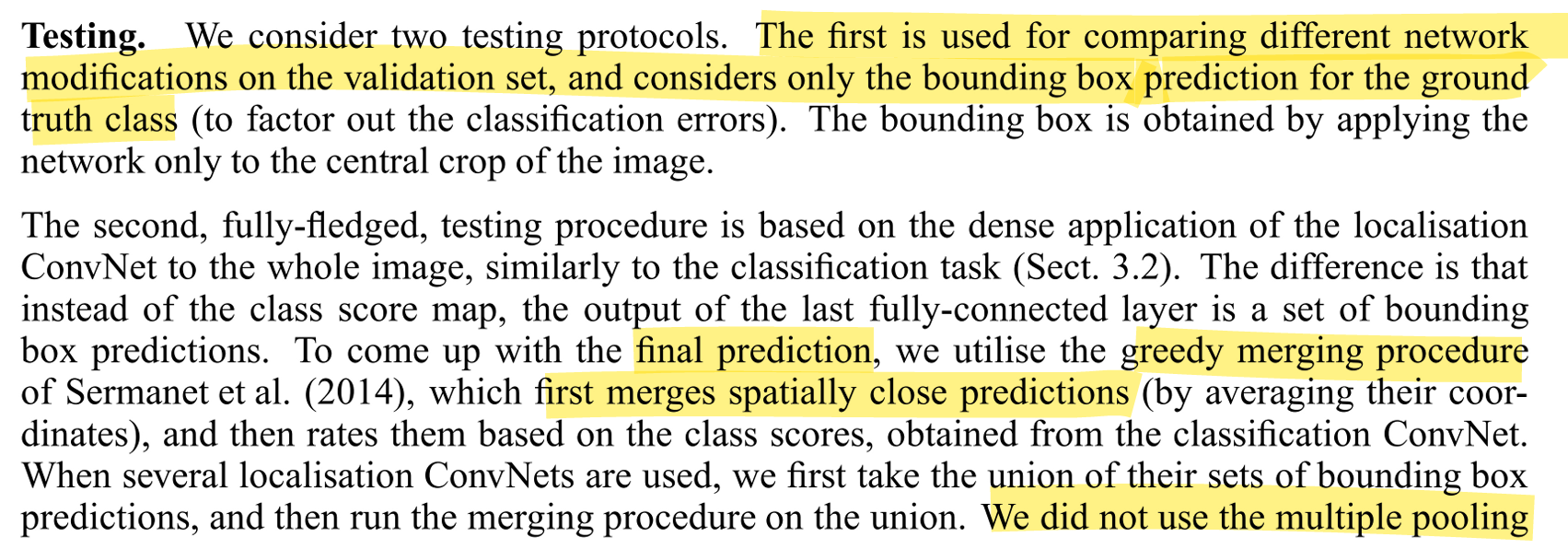
Localization :



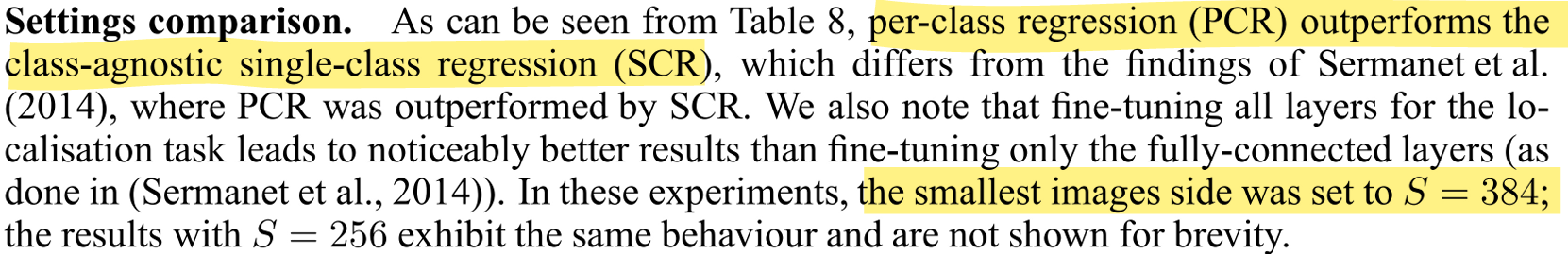
Training :



Testing :



Localization Experiments :





Comparison with the state of art in ILSVRC localization :

