VGG

Contents

- Ongoing studies
  - VGGNet

1. VGG

- A CNN network created by Oxford's Visual Geometry Group.
- As the layers get deeper, the number of parameters increases, making calculations more intensive and time-consuming. It's interesting to see how they overcame this challenge.

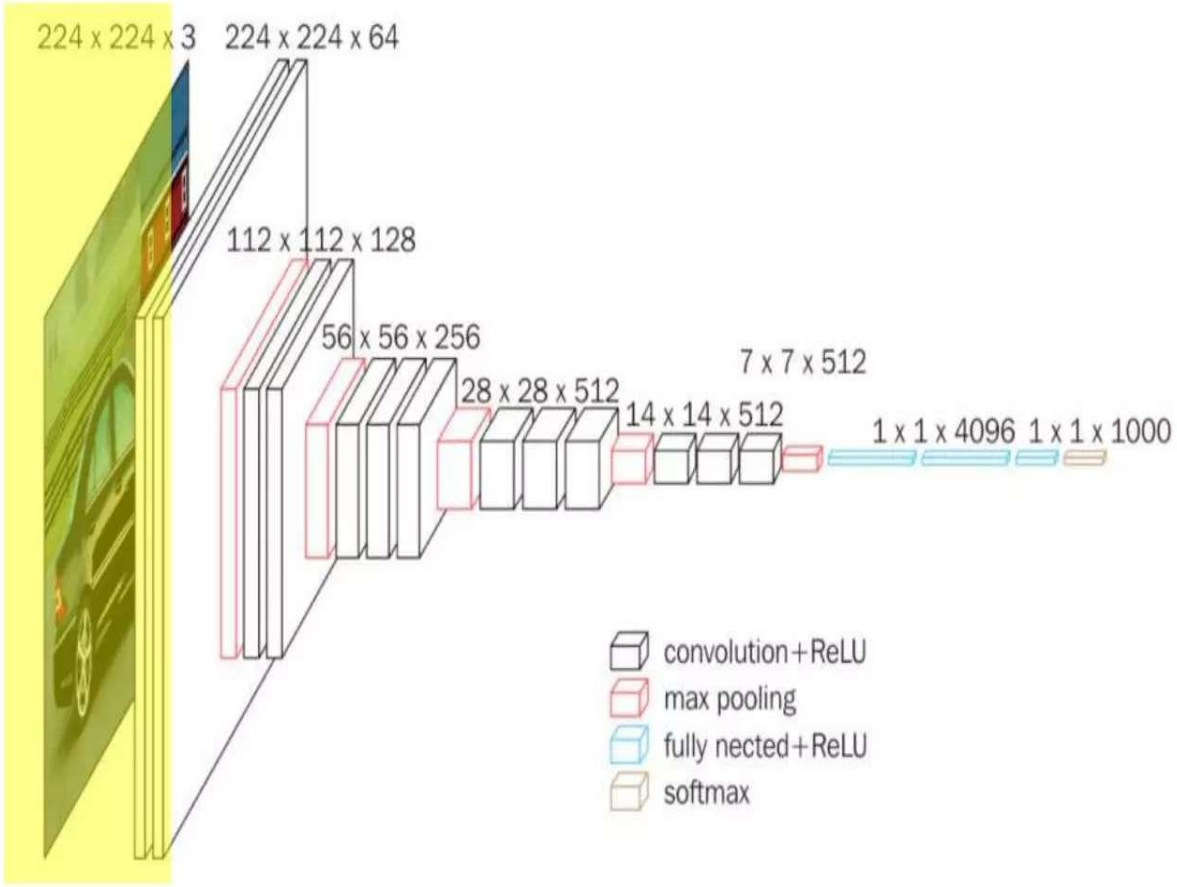
1. VGG - introduction

- The author focused on deepening the depth to improve performance.
- As a result, not only on the ImageNet dataset but also on various other image datasets, they demonstrated performance comparable to the state of the art (SOTA).

1. VGG - Configurations

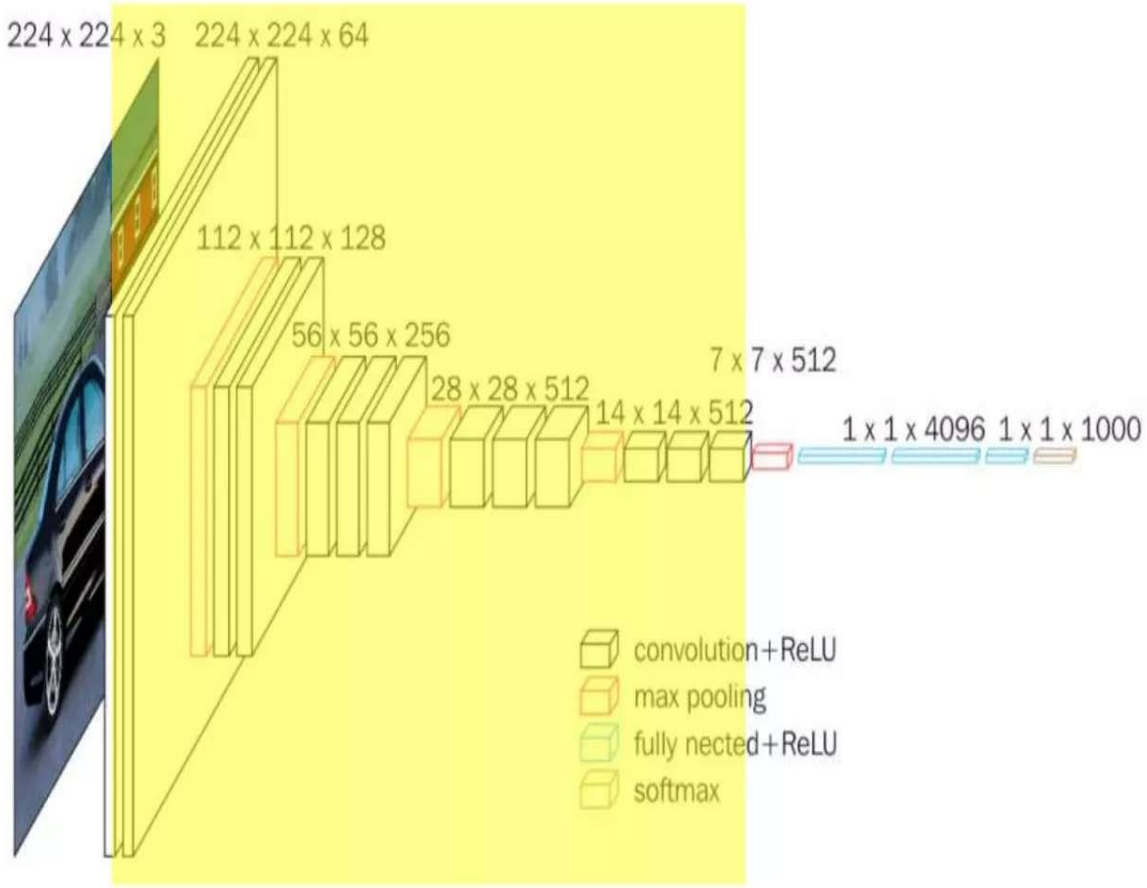
- To measure the improved performance as the depth of the ConvNet increases, all ConvNet layer configurations were designed based on the same principle.
- First, we'll examine the general layout of the ConvNet, and then delve into the details of the specific configurations used for evaluation.
- After that, we'll compare it with the previous state of the art (SOTA) and discuss the findings.

1. VGG - Architecture



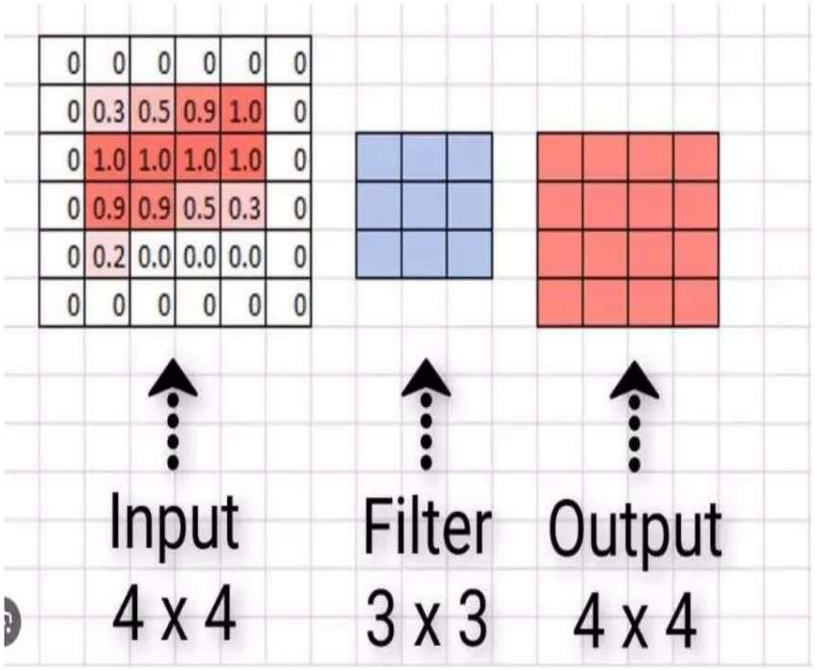
- During the training process, an RGB image of size 224x224 is provided as input.
- The only preprocessing we do is subtracting the mean RGB value, computed on the training set, from each pixel

1. VGG - Architecture



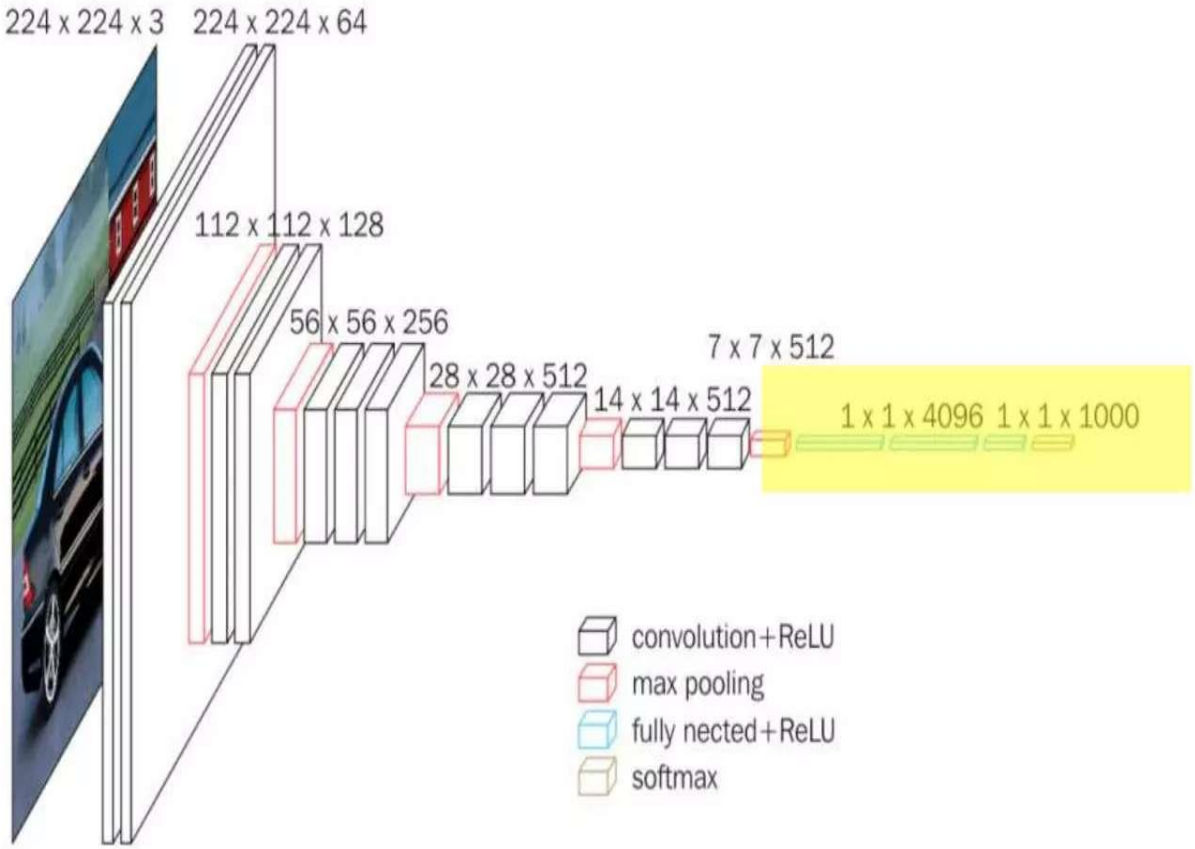
- After preprocessing, the image passes through several ConvNet layers, and during this process, it goes through kernels of very small sizes:
  - 3x3 kernel: This is the minimum kernel size to consider the adjacency in all four directions (up, down, left,

1. VGG - Architecture



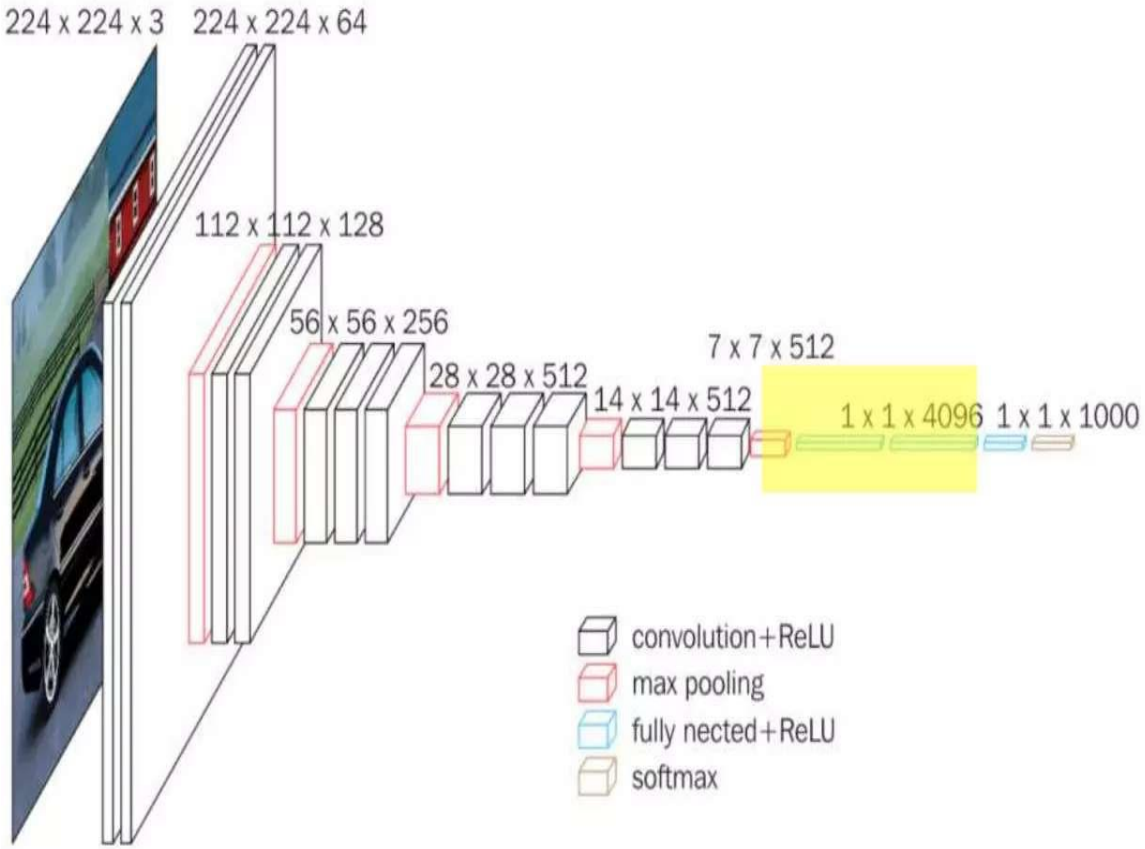
- In the Conv layer, zero-padding is set to 1 to ensure the size of the previous input size.(3x3 kernel)

1. VGG - Architecture



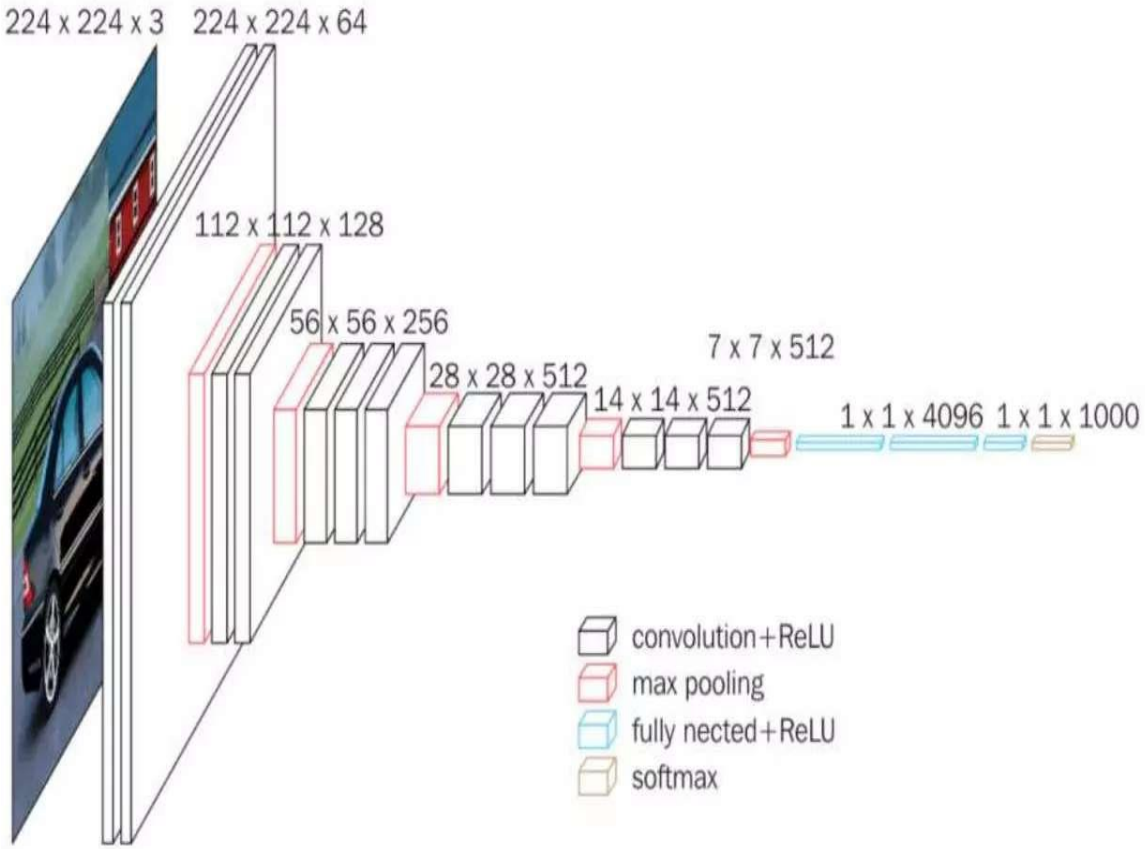
- Following the Conv layers, there are three Fully Connected (FC) layers. The first two layers each have 4096 channels, and the last one consists of 1000 channels.
- The final layer is a softmax layer designed to classify into 1000 classes.

1. VGG - Architecture



- The training was regularised by weight decay (the L2 penalty multiplier set to 5 X10^−4) and dropout regularisation for the first two fully-connected layers (dropout ratio set to 0.5).

1. VGG - Architecture



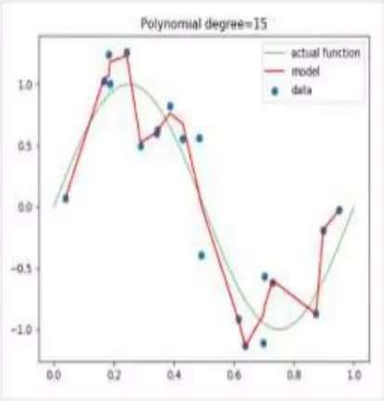
- The ReLU non-linear function is used in all hidden layers.
- They argued, contrary to AlexNet, that LRN (Local Response Normalization) doesn't help improve performance. Instead, it only increases memory usage and computation time, so they didn't use it.

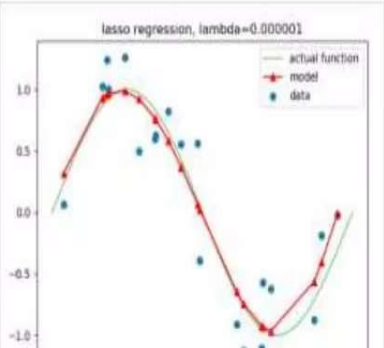
1. VGG - DISCUSSION

- While other models used sizes like 11x11 or 7x7, the used much smaller kernel like 3x3 and 1x1 sizes.
- There are two reasons for this choice:
  - To make the decision function more discriminative.
  - To reduce the number of parameters.

Weight decay in VGG

- Weight decay
- One of the ways to solve the overfitting problem.
- To prevent weight from taking on too large a value, put a penalty in the loss function when weight gets too large(L1 Regularization, L2 Regularizatio)
- apply weight decay, can get away with overfitting as shown in the second image above





- End. QnA