RepVGG: Making VGG-style ConvNets Great Again

Felix Hausberger
Universität Heidelberg
Grabengasse 1, 69117 Heidelberg
eb260@stud.uni-heidelberg.de

Abstract

1. Introduction

2. Related Work

The VGG architecture was introduced in [3]. One of its key finings was to prefer deep CNNs (16-19 weight layers) with small receptive fields induced by using small kernels over shallow CNNs with bigger receptive fields. Therefore a configuration of 3x3 kernels with stride 1 were used. This not only helps to strengthen the discriminative character of the network as the non-linear activation function (ReLU) is applied more often but also keeps the number of parameters to train lower. To increase the non-linearity without affecting the related receptive fields even 1x1 kernels were considered in more deeper architectures. Only by using simple convolutional, max-pooling and fully connected layers at the end of the network, VGG achieved a 24.4 top-1 validation error score during ILSVRC-2014 (single net performance). [3]

Regarding the top-5 validation error score VGG got beaten by GoogLeNet with 6.67 compared to 7.32 from VGG. GoogLeNet uses a very deep CNN with 22 trainable layers with nine of them being the novel inception modules. To counter the higher computational costs that come with deeper architectures and also to prevent overfitting when having a limited dataset, an inception module uses 1x1 kernels for dimension reduction. To better recognize objects at various scales, an inception modules applies 1x1, 3x3 and 5x5 kernels simultaniously and bundles its results for the next layer making the network architecture also wider than others. [1]

ResNet [2].

- structural re-parameterization technique - ResNet-50, - ResNet-101 - DenseNet - EfficientNet - RegNet - automatic, or manual architecture search - search compound scaling strategy

3. Approach

4. Experiments

5. Discussion

6. Conclusion

References

[1] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. Going deeper with convolutions, 2014. 0

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep residual learning for image recognition, 2015. 0

[3] Andrew Zisserman Karen Simonyan. Very deep convolutional networks for large-scale image recognition, 2014. 0