Approximated Oracle Filter Pruning for Destructive CNN Width Optimization
Appendix

Optimized Structures Discovered by AOFP

VGG on CIFAR-10.

In the re-designed VGG structure, each layer has 44, 80, 160, 180, 360, 360, 256, 224, 192, 56, 80, 192, 192 filters, respectively. Compared to the baseline, this model requires a roughly equal amount of FLOPs (312 v.s. 313 MFLOPs), but runs $1.27\times$ as fast (8099 examples/sec v.s. 6366 examples/sec), using CUDA9.2 and Tensorflow 1.10 on a GTX 1080Ti GPU. Interestingly, the wide layers at the early stages do not cause computational bottlenecks.

ResNet-50 on ImageNet.

The discovered structure of the re-designed ResNet-50 is depicted in Fig. 1.

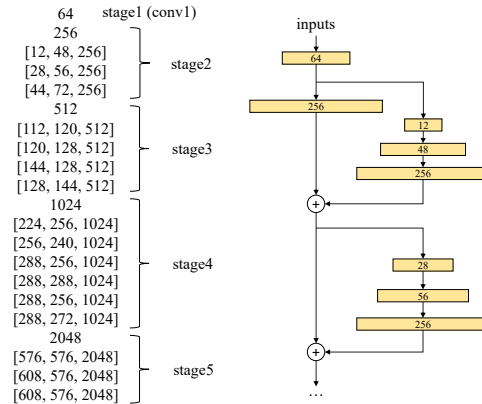A comparison of the re-designed structures and the original models are presented in Fig. 2.

Figure 1. The structure of the re-designed ResNet-50. We present the number of filters of each convolutional on the left, and depict several layers at the beginning for example on the right.
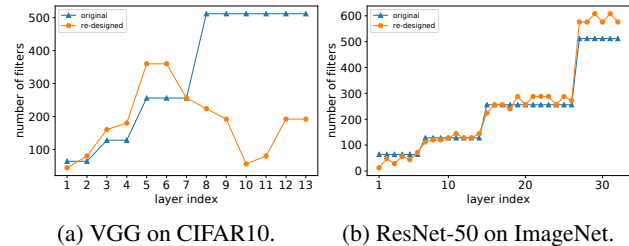

(a) VGG on CIFAR10.   (b) ResNet-50 on ImageNet.

Figure 2. Layer width of the re-designed models in comparison with the original. Note again that only the internal layers of ResNet-50 (i.e., the first two layers in each residual block) are shown.