Learning Filter Basis for Convolutional Neural Network Compression: Supplementary Material

Yawei Li[1*], Shuhang Gu[1*], Luc Van Gool[1,2], Radu Timofte[1]
[1]Computer Vision Lab, ETH Zurich, Switzerland, [2]KU Leuven, Belgium
{yawei.li, shuhang.gu, vangool, radu.timofte}@vision.ee.ethz.ch

Number of basis m	Channel c	Transition Layer Split
24	444	1
26	444	1
32	444	1
36	444	6
38	444	12

Table 1. DenseNet-12-40 compression configuration for Table 4 in the main paper. The basis set is shared by all of the DenseBlocks. For lower layer DenseBlocks, a slice of the shared basis is used as the basis of that layer. For the former three configurations, we do not compress the transition layers in DenseNet. But for the latter two, the transition layers are also compressed with the specified number of splits.

1. Compressed Network Configuration

The basis configurations of our filter basis learning method for different networks including DenseNet [2], ResNet [1], VGG [6], EDSR [4], EDSR-8-128, SRResNet [3] are shown in Table 1, Table 2, and Table 3. For DenseNet, we used the network-wise basis sharing. For ResNet, we used group-wise basis sharing. We also tried basis sharing within the residual block for EDSR.

We reimplemented the network compression method Factor [7] and Group [5]. For the Factor method, to compare the methods fairly, we use two and three single intra-channel convolutional layers (SIC layer) [7] in Table 2, two SIC layers in Table 5 and Table 6, and one SIC layer in Table 3 to substitute one standard convolutional layer. To keep the number of parameter of the Group method [5] at the same level with other methods, the group size is set to 8, and 64 to approximate ResNet, and VGG respectively. To compress DenseNet, 3 groups are used for the first 20 DenseBlocks while 6 groups are used for the rest DenseBlocks.

*Equal contribution

Residual Block Group	Number of basis m	Channel c
Group One	24	16
Group Two	48	32
Group Three	84	64

Table 2. ResNet-56 compression configuration. There are 27 residual blocks in ResNet-56, distributed into three groups with increasing number of channels but reducing resolution. The basis is shared by the convolutions within the same group. This configuration corresponds to the ResNet-56 entry in Table 5 of the main paper.

Network	Number of basis m	Channel c
SRResNet (Basis-64-14)	14	64
SRResNet (Basis-32-32)	32	32
EDSR-8-128 (Basis-128-27)	27	128
EDSR-8-128 (Basis-128-40)	40	128
EDSR (Basis)	32	256
EDSR (Basis-S)	32	256
VGG-16	128	128

Table 3. Compression configuration of SRResNet, EDSR-8-128, EDSR, and VGG-16. 'Basis' means that there is a unique basis for each convolutional layer. 'Basis-S' means that the basis is shared by the two convolutional layers within the residual block. For VGG-16, the first three convolutional layers are not compressed.

2. Training and Testing Error Curves for Image Classification

The error curves during training and testing for DenseNet-12-40, ResNet-56, and VGG-16 on CIFAR10 are shown in Fig. 1, Fig 3, and Fig 2, respectively. Our method shoots the lowest stable error rate for all the three networks during training and testing.

3. More Visual Results for Super-Resolution

More visual results for image super-resolution are shown in Fig. 5 and Fig. 4 for compressing SRResNet and EDSR-8-128 respectively. Compared with Factor [7] and Group [5], the SR images from our compressed model are

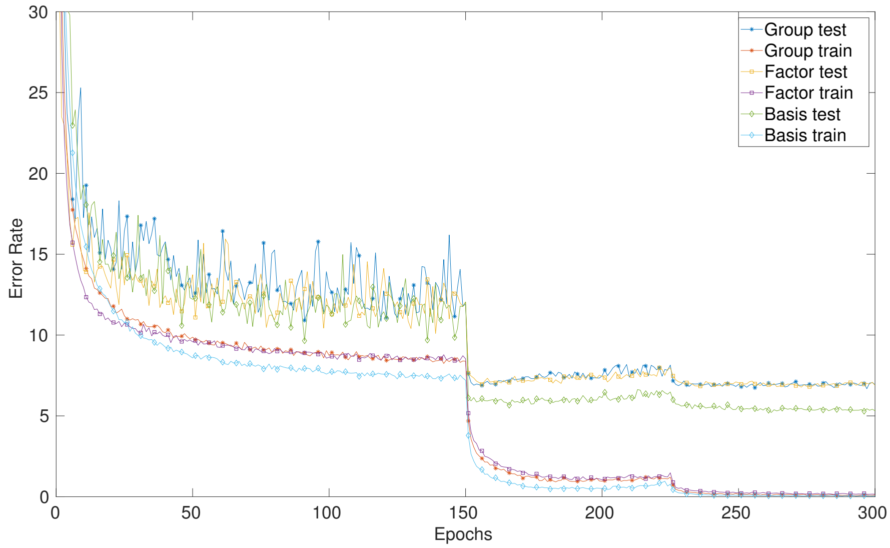

Figure 1. Training and testing error of different compression method applied on DenseNet-12-40.

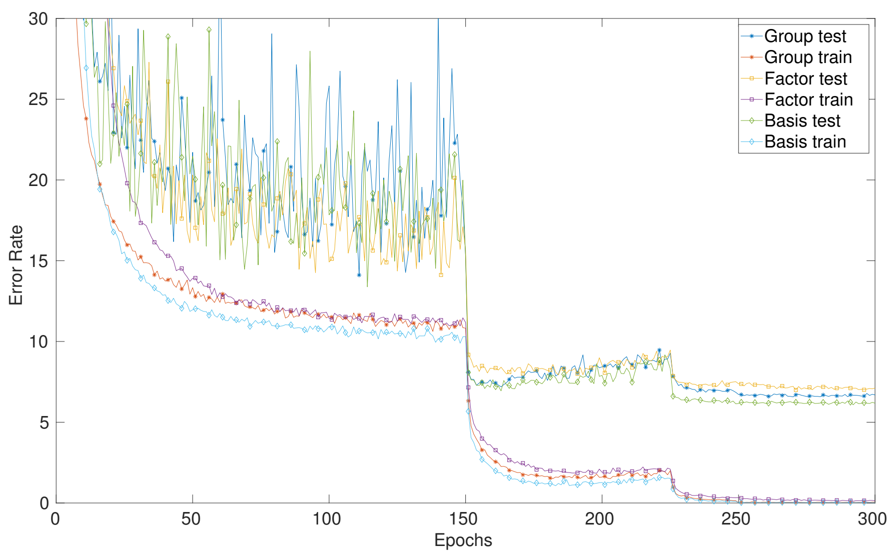

Figure 2. Training and testing error of different compression method applied on VGG-16.

very close to the baseline in terms of both visual quality and PSNR values.

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proc. CVPR*, pages 770–778, 2016. 1

[2] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proc. CVPR*, pages 2261–2269, 2017. 1

[3] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proc. CVPR*, pages 105–114, 2017. 1

[4] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proc. CVPRW*, pages 1132–1140, 2017. 1

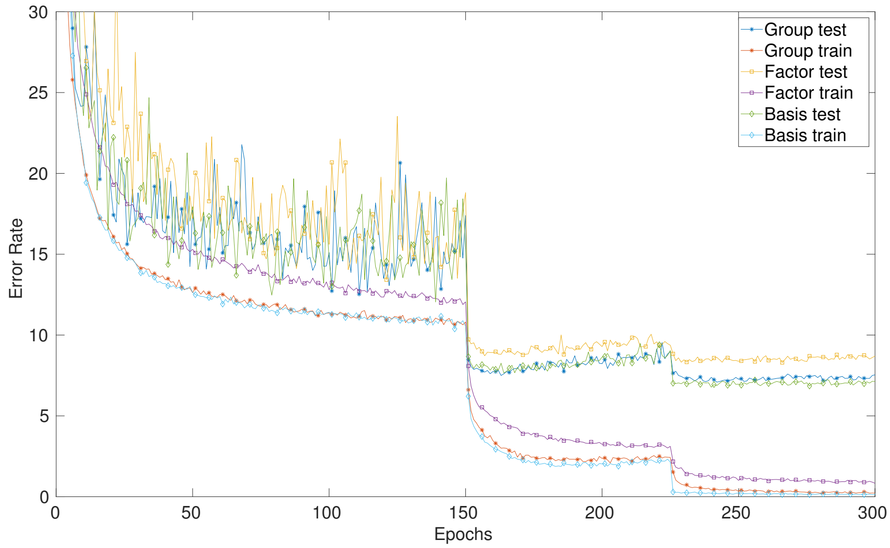

Figure 3. Training and testing error of different compression method applied on ResNet-56.

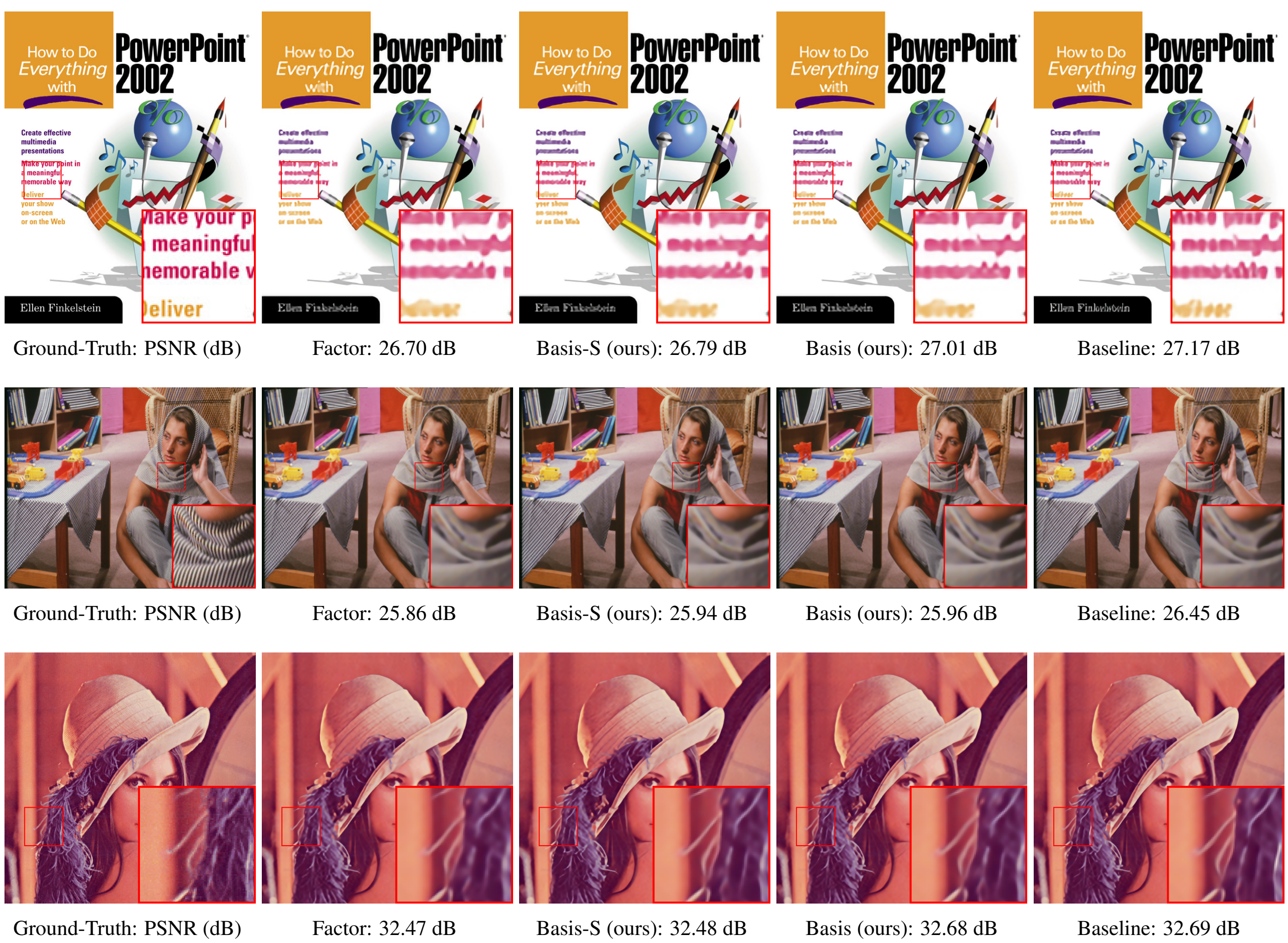

Figure 4. SR results for upscaling factor $\times 4$. Network compression methods are applied on EDSR. PSNR values are reported.

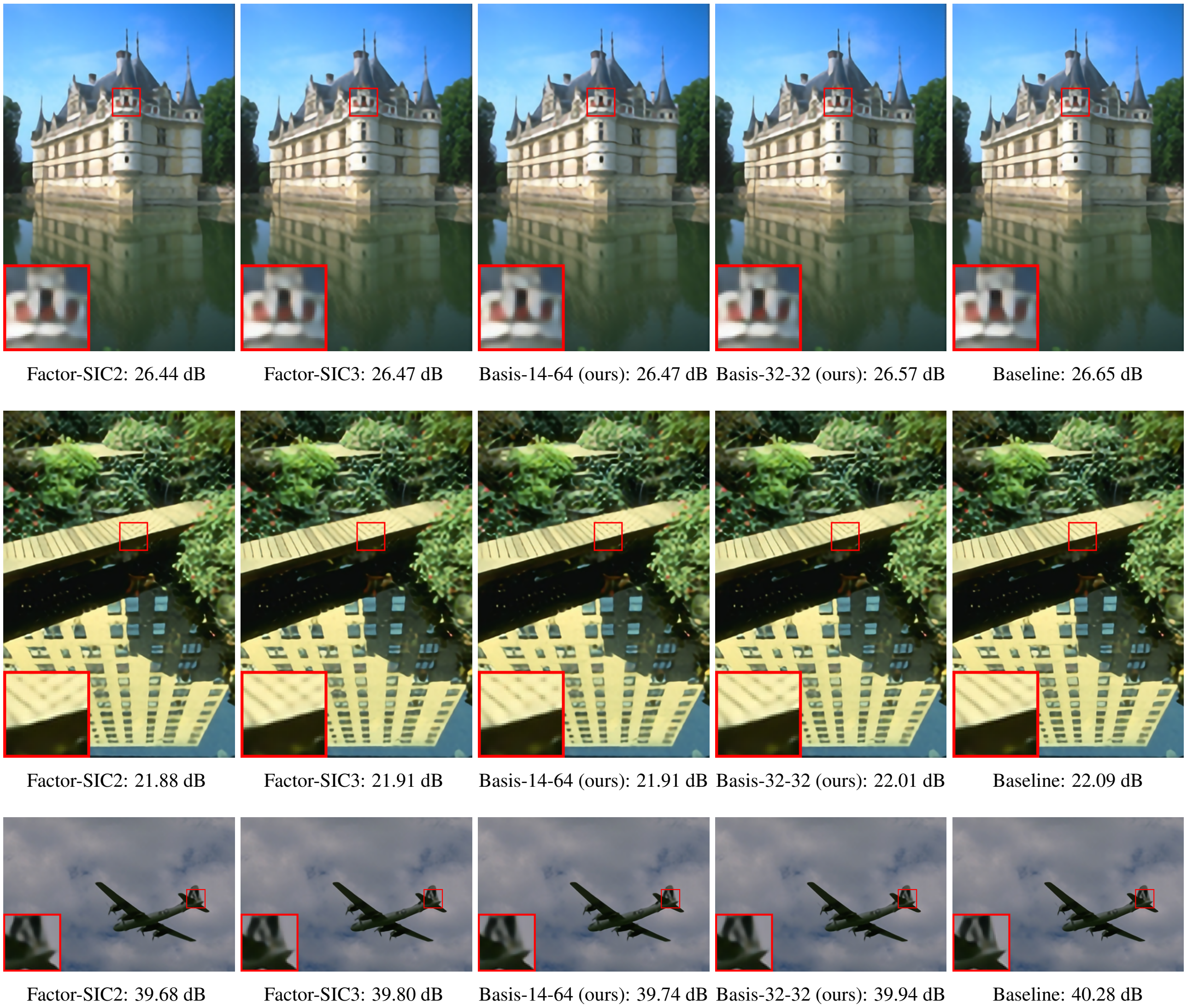

Figure 5. SR results for upscaling factor $\times 4$. Network compression methods are applied on SRResNet. PSNR values are reported.

[5] Bo Peng, Wenming Tan, Zheyang Li, Shun Zhang, Di Xie, and Shiliang Pu. Extreme network compression via filter group approximation. In *Proc. ECCV*, pages 300–316, 2018. 1

[6] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1

[7] Min Wang, Baoyuan Liu, and Hassan Foroosh. Factorized convolutional neural networks. In *Proc. ICCV*, pages 545–553, 2017. 1