NTIRE Workshop
CVPR 2021

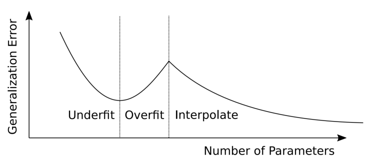
Overparametrization of HyperNetworks Enables Fast Neural Image Enhancement

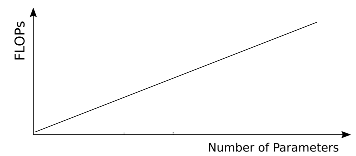
Lorenz K. Muller, Computing Systems Laboratory, Zurich Research Center, Huawei Technologies Switzerland
lorenz.mueller@huawei.com, link to full paper

HUAWEI

Problem

Deep ConvNets excel at tasks like demoisaicing, denoising and super-resolution. For mobile devices these networks often need too many FLOPs. A ConvNet with fewer FLOPs, also has few parameters. This could be a problem, because neural networks with very many parameters usually generalize best.

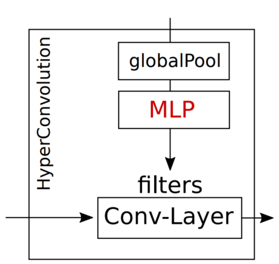




How can we reconcile high quality with few FLOPs in a ConvNet?

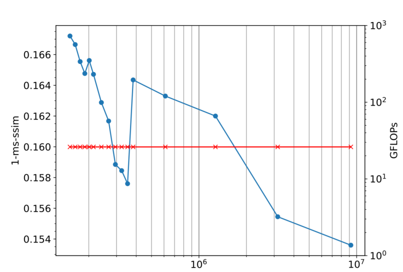
Method

We use a HyperNetwork [4,5] to predict the filters of convolution layers, similar to [6]. We design this HyperNetwork such that it allows adding parameters at constant FLOP count in the large image limit. For this we define the 'HyperConvolution' (HC) below and resize the MLP inside it.



The HC-block takes two sets of feature maps as inputs. The first set is used to predict filters, the second set is convolved with these predicted filters.
The HC-block has some favorable properties:
1) It acts at full resolution
2) It is translation-equivariant
3) It makes use of non-local information
4) Its filters are not limited to a discrete set
5) Its FLOPs-to-parameter ratio is flexible

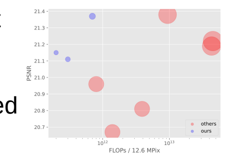
When we build a Unet-like network with such blocks, we observe a generalization curve consistent with a 'double-descent' shape as we scale up the MLPs in the HC-blocks. Note that the FLOP count stays nearly constant, while we move into the favorable overparameterized regime.



Experiments

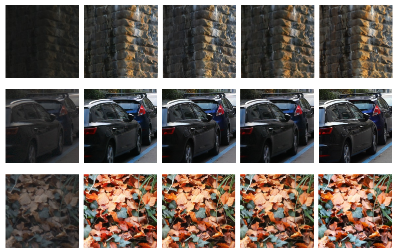
Zurich Raw to DSLR (ZRR)

We evaluate this approach on ZRR [1]. We find that our approach substantially improves MS-SSIM and SSIM, matches PSNR and reduces computational cost (>10x fewer FLOPs and memory use) compared to SOTA single network performance.



Network	FLOPs	Param.s	CPU time	Conv. Mem.	PSNR	MS-SSIM	SSIM
PyNet [1]	43 T	47 M	120s	29.7 Gb	21.19	**0.8620**	-
PyNet-CA [2]	45 T	51 M	131s	32.8 Gb	21.22	0.8549	0.7360
AWNet 4-channel [3]	9.4 T	52 M	55s	27.8 Gb	**21.38**	0.8590	0.7451
Ours no HC	1.3 T	0.4 M	11s	3.8 Gb	19.93	0.8463	0.7213
Ours no HC	0.6 T	0.2 M	7s	2.6 Gb	19.82	0.8446	0.7185
Ours I	0.7 T	276 M	12s	3.4 Gb	**21.37**	**0.8640**	**0.7509**
Ours II	0.3 T	95 M	6s	2.2 Gb	21.11	0.8618	0.7466
Ours III	0.2 T	90 M	5s	1.8 Gb	21.15	0.8617	**0.7471**

Image details



RAW AWNet PyNet Ours Target

On visual inspection the image quality seems comparable to other methods. The run-time, FLOPs and memory use of our method are preferable, however.

We note that code and a pre-trained network for this task is given in the supplemental.

Full-size images

AWNet

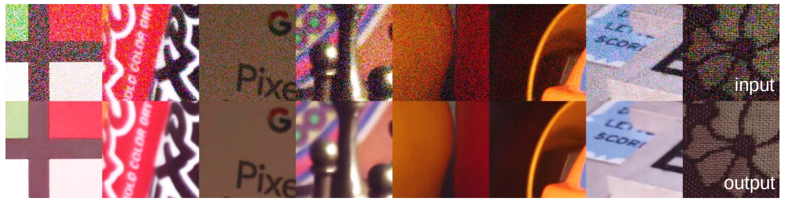
Ours

Smartphone Image Denoising Dataset (SIDD)

In a second evaluation we start from VDN [7] trained on SIDD [8] to denoise images. We fix the task, cost-function, and training details. We only modify the network architecture minimally to contain HC-blocks instead of convolution layers. We then resize the HC-blocks such that the resulting network is smaller than the original one, train it and compare the performance on SIDD (sRGB variant).
We achieve a speed-up of ca. 2x and a slight improvement in fidelity.

Network	FLOPs	Param.s	CPU time	Conv. Mem.	PSNR	SSIM
VDN [7]	9.5 T	7.8 M	3.1 s/Mpix	2.3 GB	39.26	0.955
Ours I	1.4 T	55.0 M	1.5 s/Mpix	1.0 GB	39.23	0.957
Ours II	2.9 T	119.6 M	2.5 s/Mpix	1.4 GB	**39.27**	**0.957**

Image details



Conclusions

We propose the use of HyperNetworks to break the fixed parameter to FLOPs ratio in ConvNets. We find that this yields significant speed-ups and memory reductions in neural image enhancement tasks at matched or improved fidelity. More details can be found in our paper (link).

References

[1]: Ignatov, Andrey, Luc Van Gool, and Radu Timofte. "Replacing mobile camera isp with a single deep learning model." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020
[2]: Kim, Byung-Hoon, et al. "PyNET-CA: enhanced PyNET with channel attention for end-to-end mobile image signal processing." *European Conference on Computer Vision*. Springer, Cham, 2020.
[3]: Dai, Linhui, Xiaohong Liu, Chengqi Li, and Jun Chen. "AWNet: Attentive Wavelet Network for Image ISP." *arXiv preprint arXiv:2008.09228*. 2020
[4]: Ha, David, Andrew Dai, and Quoc V. Le. "Hypernetworks." *arXiv preprint arXiv:1609.09106*. 2016
[5]: Schmidhuber, Jürgen. "Learning to control fast-weight memories: An alternative to dynamic recurrent networks." Neural Computation 4.1 (1992): 131-139
[6]: Klein, Benjamin, Lior Wolf, and Yehuda Afek. "A dynamic convolutional layer for short range weather prediction." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015
[7]: Yue, Zongsheng and Yong, Hongwei and Zhao, Qian and Meng, Deyu and Zhang, Lei. “Variational Denoising Network: Toward Blind Noise Modeling and Removal.” Advances in Neural Information Processing Systems 2019
[8]: Abdelhamed, Abdelrahman, Stephen Lin, and Michael S. Brown. "A high-quality denoising dataset for smartphone cameras." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018