FlexConv: Continuous Kernel Convolutions with Differentiable Kernel Sizes

Robert-Jan Bruintjes*, David W. Romero*
Erik J. Bekkers, Jakub M. Tomczak, Jan C. van Gemert, Mark Hoogendoorn

1

Modern CNNs

- Fixed, small kernels
- Pooling to process large images
- Adapted to specific image size

Problems

- Each dataset needs custom architecture
- Cannot adapt to new scales
- Not clear if small kernels are ideal

2

FlexConv

Free-form kernel

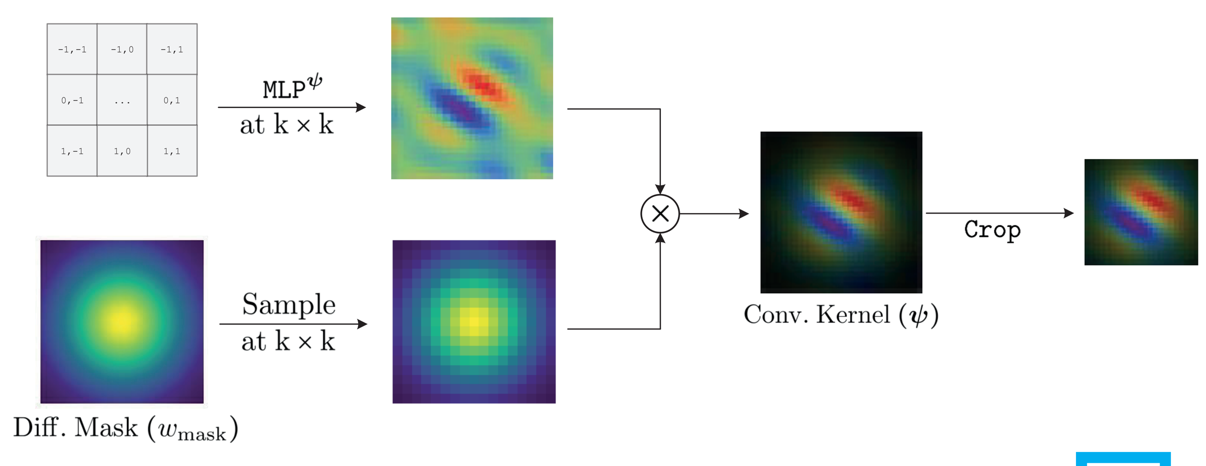
- Extension of CKConv [3] to 2D: MLP takes pixels and predicts kernel
- Constant parameters for any kernel size
- New MLP: "MAGNet"

Learnable kernel size

- Added learnable Gaussian mask
- Crop where zero (2 stddevs) for faster convolution

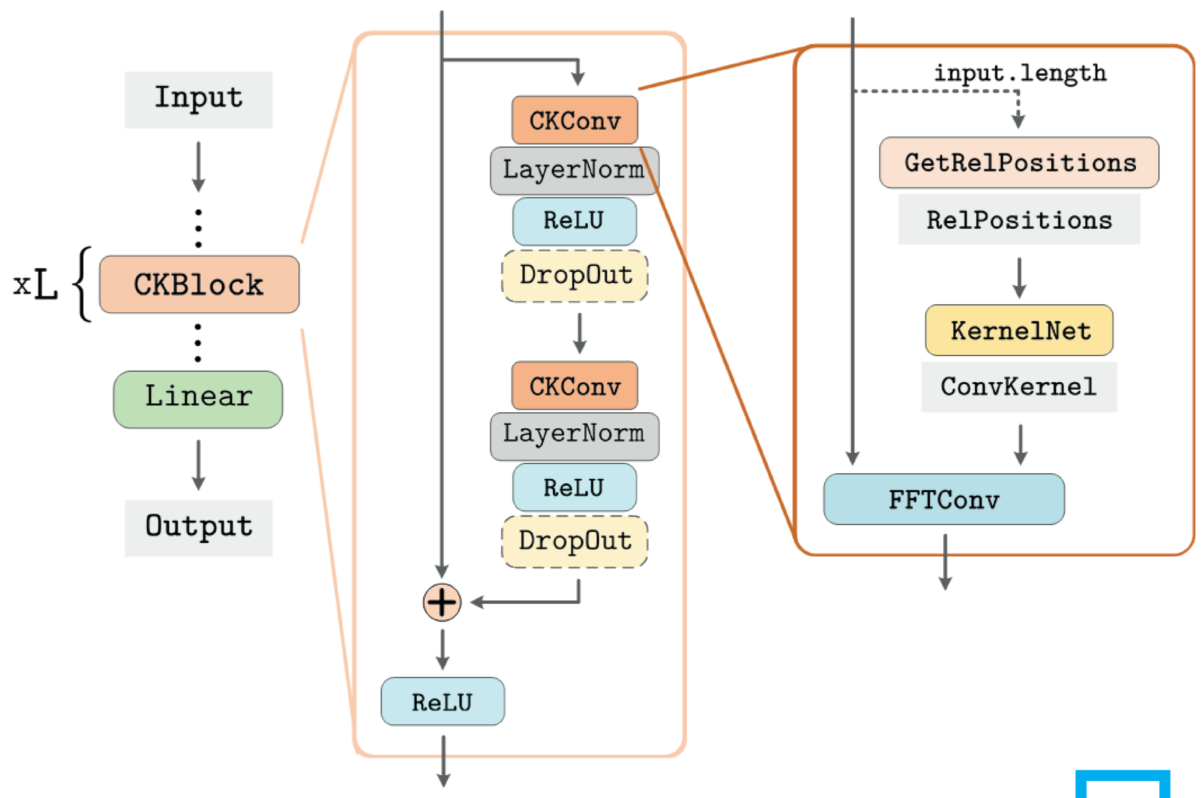
Works at any resolution

- MAGNet allows for frequency analysis
- We regularize highest frequency to prevent aliasing when upscaling



1

FlexConv



2

FlexNet

4

Related Works

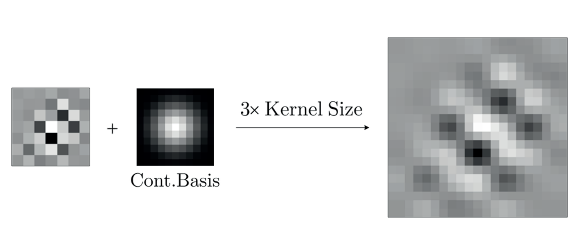
1. Dilation of normal kernels [5] or filter bases [2, 4, 6] (Fig. 3a)
2. Deformable kernels [1] (Fig. 3b)

Limited frequency response, unlike FlexConv (Fig. 3c)

3

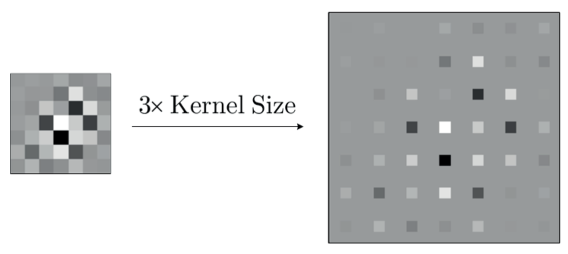
FlexNet

- Residual FlexConv blocks
- No pooling layers
- Adapts to image size



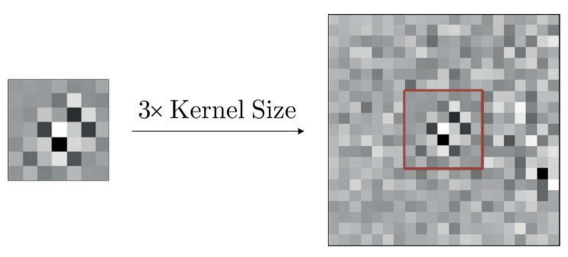
3a

Dilation [5]



3b

Deformable [1]



3c

Ours

Model	Size	CIFAR-10 Acc.
CIFARResNet-32 (He et al., 2016)	0.53M	92.5*
N-Jet-ALLCNN (Pintea et al., 2021)	1.07M	92.5 ± 0.1*
N-Jet-Resnet32 (Pintea et al., 2021)	0.52M	92.3 ± 0.3*
DCN-σ^{ji} (Tomen et al., 2021)	0.47M	89.7 ± 0.3*
FlexNet-7 w/ convolutions (k = 3)	0.17M	89.5 ± 0.3
FlexNet-7 w/ convolutions (k = 33)	20.0M	78.0 ± 0.3
FlexNet-7 w/ N-Jets (Pintea et al., 2021)	0.70M	91.7 ± 0.1
FlexNet-3	0.27M	91.3
FlexNet-5	0.44M	92.5
FlexNet-7	0.67M	**92.7**

Exp: 2D datasets

Model	Size	sMNIST	pMNIST	sCIFAR10	npCIFAR10
DilRNN (Chang et al., 2017)	44K	98.0	96.1	-	-
IndRNN (Li et al., 2018)	83K	99.0	96.0	-	-
TCN (Bai et al., 2018a)	70K	99.0	97.2	-	-
r-LSTM (Trinh et al., 2018)	0.5M	98.4	95.2	72.2	-
Self-Att. (Trinh et al., 2018)	0.5M	98.9	97.9	62.2	-
TrellisNet (Bai et al., 2018b)	8M	99.20	98.13	73.42	-
URLSTM (Gu et al., 2020b)	-	99.28	96.96	71.00	-
URGRU + Zoneout (Gu et al., 2020b)	-	99.27	96.51	**74.40**	-
HiPPO (Gu et al., 2020a)	0.5M	-	**98.30**	-	-
Lipschitz RNN (Erichson et al., 2020)	158K	99.4	97.3	64.2	59.0
coRNN (Rusch & Mishra, 2020)	134K	**99.4**	97.3	-	59.0
UnICORNN (Rusch & Mishra, 2021)	135K	-	98.4	-	**62.4**
pLMU (Chilkuri & Eliasmith, 2021)	165K	-	98.49	-	-
CKTCN (2-Blocks)	98K	99.31	98.00	62.25	
CKTCN-Big (2-Blocks)	1M	99.32	**98.54**	63.74	
CKTCN$_{\text{MAGNet}}$ (2-Blocks)	105K				
FlexTCN (2-Blocks)	108K	**99.60**	**98.61**	**78.99**	**67.11**
FlexTCN (4-Blocks)	241K	**99.60**	**98.72**	**80.26**	**67.42**
FlexTCN (6-Blocks)	375K	**99.62**	**98.63**	**80.82**	**69.87**

Exp: 1D datasets

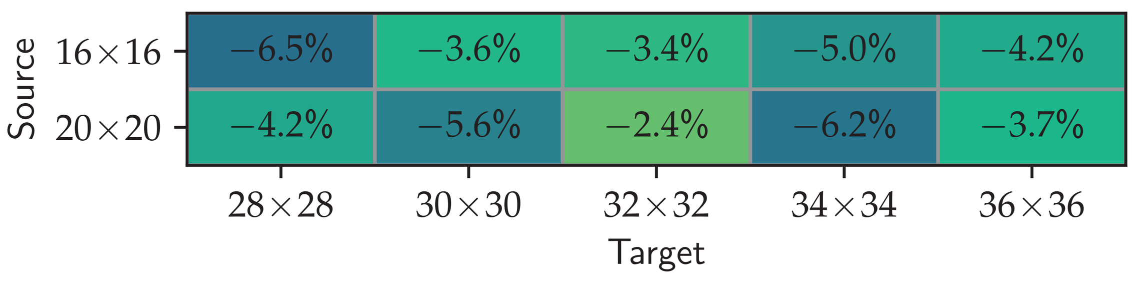
4

5

Experiments

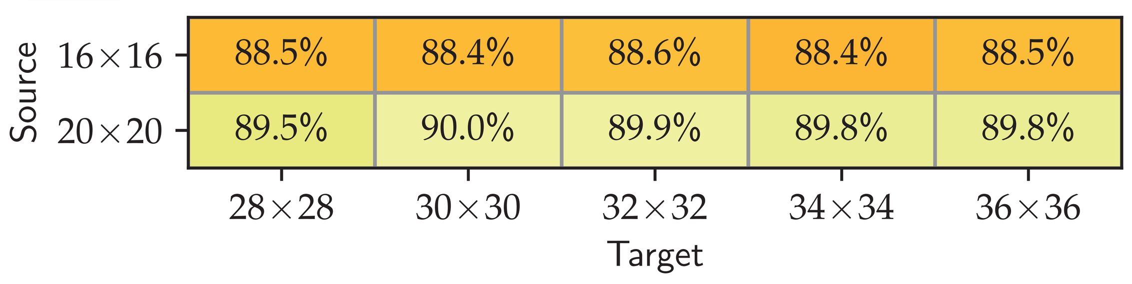
- 2D: ***competitive*** with ResNet, outperforms filter bases works (Fig. 4)
- 1D: MAGNet is ***state-of-the-art*** (Fig. 5)
- ***Minimal*** precision loss with aliasing regularization

5



Exp: CIFAR-10, acc. at new resolution

6a



Exp: CIFAR-10, acc. after fine-tuning

6b

6

Learned networks

- CIFAR-10: small kernels in early layers, large kernels in later layers

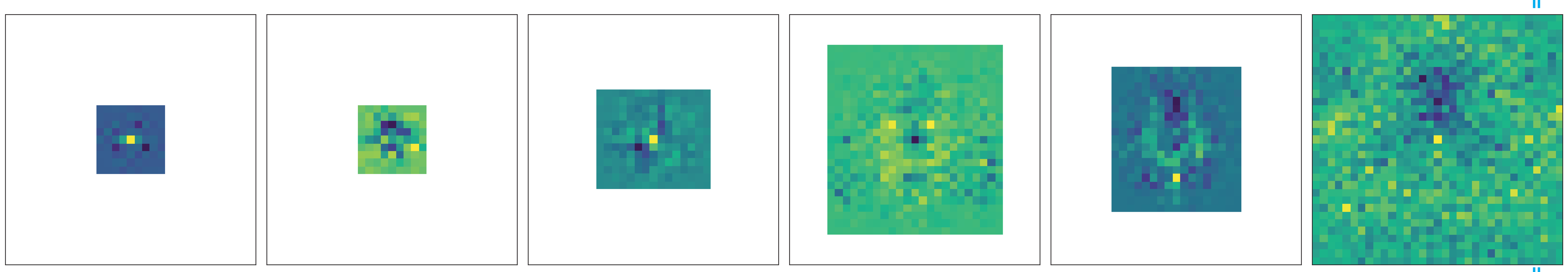
7

Conclusions

- FlexConv has unrestricted freq. band, where other works have limited bands, tied to kernel size
- FlexNets can be used on any dataset, without tuning kernel size or pooling
- FlexNets beat SOTA 1D datasets, and beat filter bases works on 2D
- Shallow nets with global kernel size can be competitive with deep, local kernels

Limitations

- More parameters for small kernels than conventional convolutions
- FlexConv is too expensive to run in deep modern architectures
- Large conv kernel are still expensive to run, if FlexNet learns them



7

Kernels learned on CIFAR-10 (left to right, shallow to deep)

References

1. Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 764–773, 2017.
2. Silvia L Pintea, Nergis Tomen, Stanley F Goes, Marco Loog, and Jan C van Gemert. Resolution learning in deep convolutional networks using scale-space theory. *arXiv preprint arXiv:2106.03412*, 2021.
3. David W Romero, Anna Kuzina, Erik J Bekkers, Jakub M Tomczak, and Mark Hoogendoorn. CKConv: Continuous kernel convolution for sequential data. *arXiv preprint arXiv:2102.02611*, 2021.
4. Nikhil Saldanha. Frequency Learning for Structured CNN Filters with Gaussian Fractional Derivatives. Master's thesis, Delft University of Technology, 2021.
5. Evan Shelhamer, Dequan Wang, and Trevor Darrell. Blurring the line between structure and learning to optimize and adapt receptive fields. *ArXiv*, abs/1904.11487, 2019.
6. Nergis Tomen, Silvia Laura Pintea, and Jan van Gemert. Deep continuous networks. In *International Conference on Machine Learning (ICLR),* 2021.

Robert-Jan Bruintjes, TU Delft, Netherlands
rjbruin.github.io @rjbruin

TU Delft