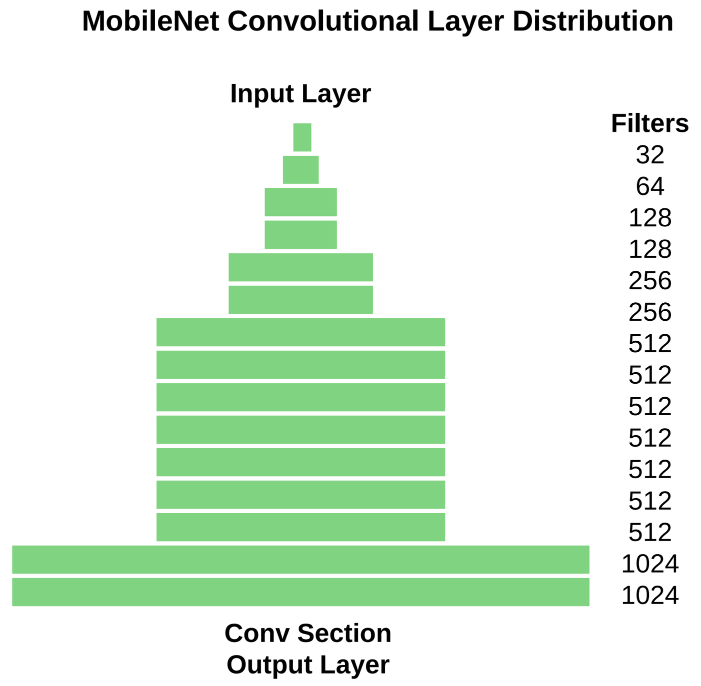
The Myth of the Pyramid

challenging the convolutional network pyramidal design

Ramon Izquierdo-Cordova, Walterio Mayol-Cuevas
Department of Computer Science, University of Bristol, United Kingdom
izquierdocr@outlook.com, walterio.mayol-cuevas@bristol.ac.uk

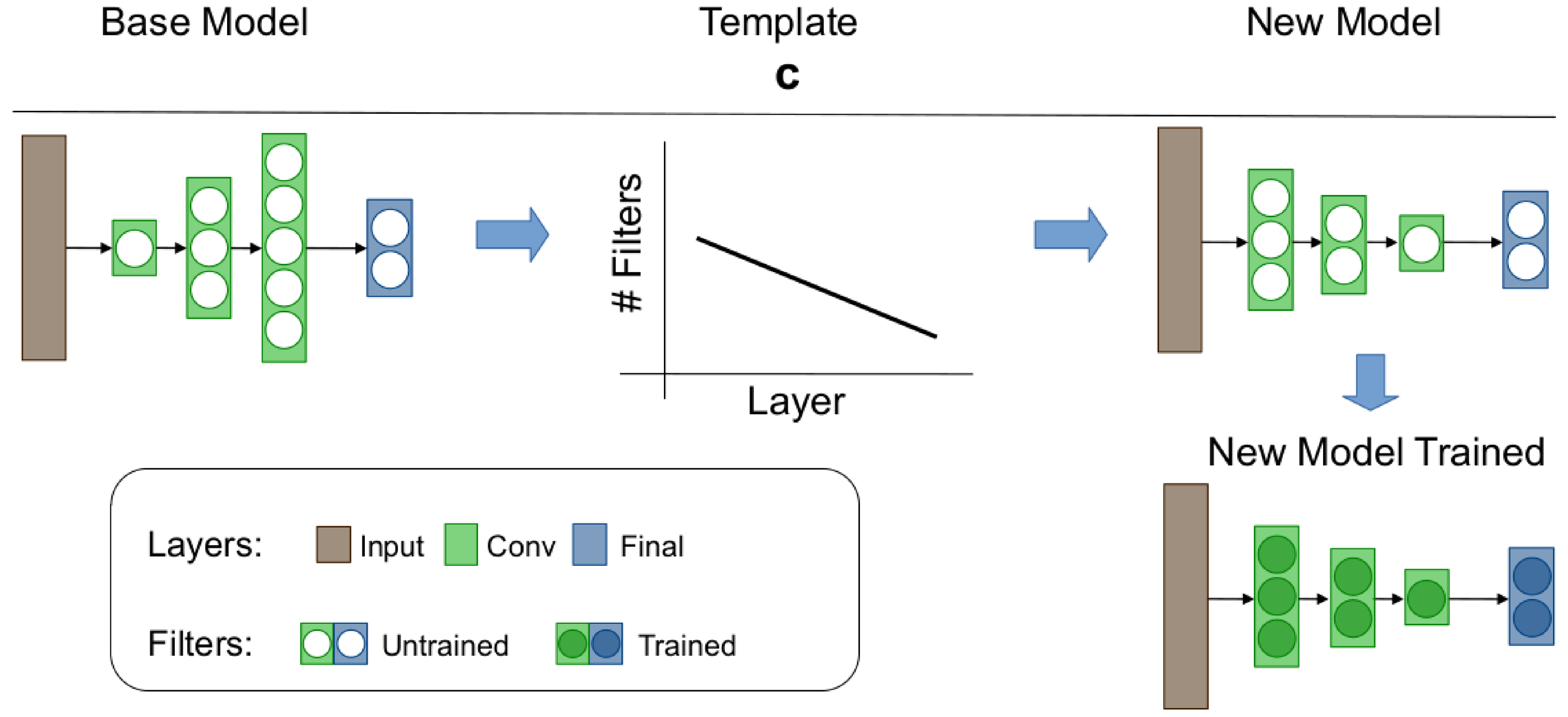
The pyramidal design

Visit Comalcalco, Tabasco. Mexico
First author home town



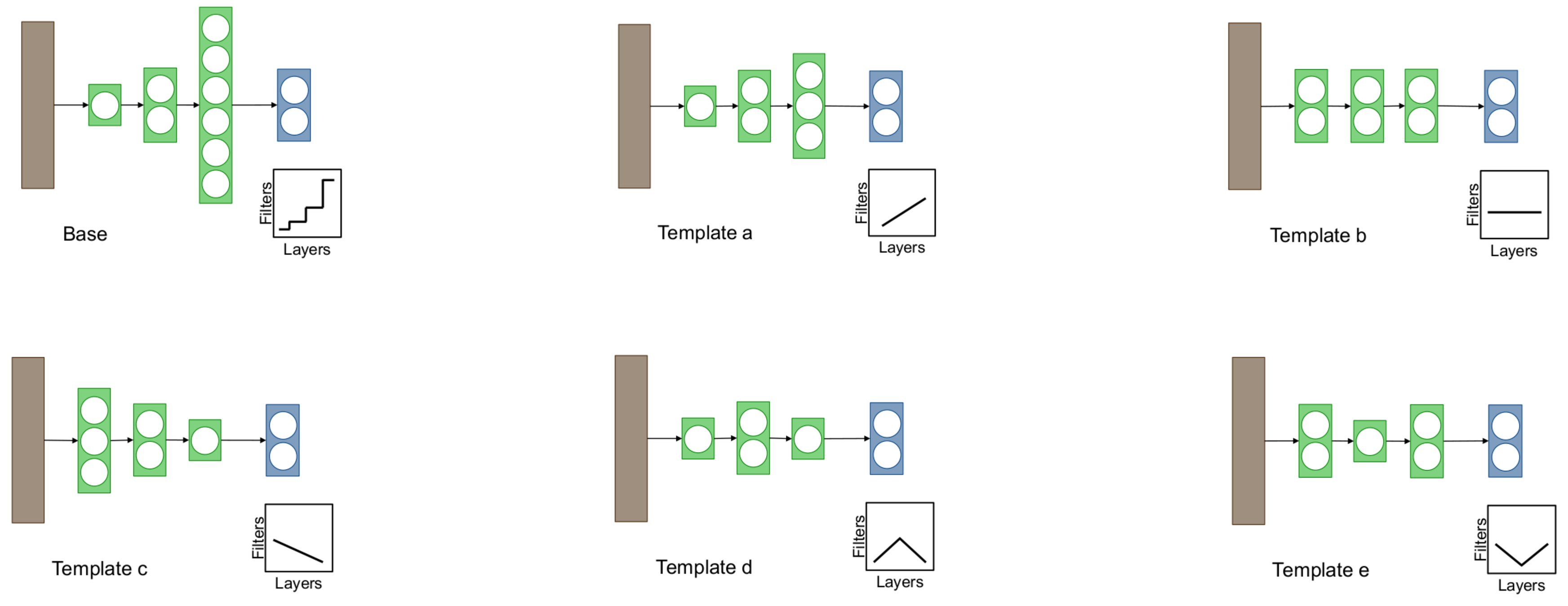
- Pyramidal CNN structure: increasing number of filters while decreasing feature map size in deeper layers.
- VGG, ResNet, GoogleNet, MobileNet, NASNet and many others CNN models follow a similar filter distribution.
- It is generally believed, yet not necessarily proven, that the pyramidal distribution of filters keeps the "richness" of the representation [1] and "controls" computational complexity [2].

Producing models with a new filter distribution



Choosing a set of basic templates

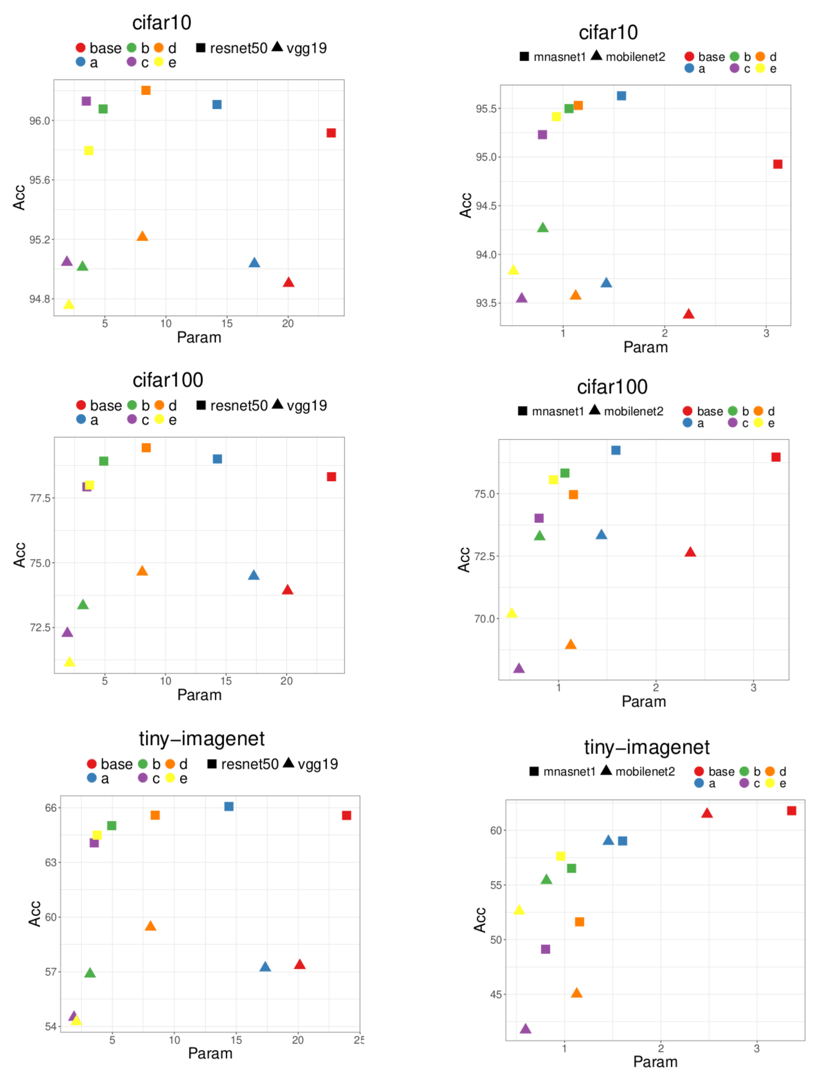
- **Keep the set small**
  Small differences between distributions lead to negligible differences in accuracy and resource consumption.
- **Keep the functions simple**
  Straight-forward linear and piece-wise linear functions easily implemented for most of the existing CNNs.
- **Reuse emergent patterns**
  Inspired by results in pruning and automatic model discovery methods [3,4,5,6]



How to fairly compare CNN models?

- Accuracy, FLOPs, memory, parameters, layers, filters, training time, inference time.
- On CIFAR datasets, VGG: 20.03 million parameters and 399 MFLOPs. ResNet: 23.52 million parameters and 1307 MFLOPs. More than three times VGG FLOPs!
- Proposal: match one metric (FLOPs) while comparing a second one (accuracy, parameters, memory footprint, inference time).

Results



Imagenet

Models		base	Templates a	b	c	d	e
VGG19	Acc	72.936	72.804	66.406	62.198	68.248	60.874
(19.6 GFLOPs)	Param	143.672	155.523	54.759	35.618	41.798	58.070
ResNet50	Acc	75.309	74.603	70.667	68.425	72.823	69.160
(4.1 GFLOPs)	Param	25.557	15.847	5.405	3.844	8.998	4.214

Audio

Metric	base	Filter Templates a	b	c	d	e
Accuracy GTZAN	85.59	85.92	87.26	85.75	**87.43**	85.08
Accuracy ESC-50	69.66	**70.33**	68.16	65.75	69.91	67.25
Param (Millions)	23.61	14.23	4.88	**3.50**	8.39	3.71
Memory (MB)	395.51	350.54	383.10	385.15	**337.40**	392.29
Inference (ms)	5.47	7.47	4.56	7.75	4.66	**4.48**

NASBench101-CIFAR10

Models		base	Templates a	b	c	d	e
Best architecture	Acc	95.35	95.20	95.02	**95.44**	95.06	95.26
3664 MFLOPs	Param	32.42	27.49	**6.44**	10.38	17.26	8.73
ResNet-like	Acc	92.64	**93.85**	91.80	92.65	91.81	92.67
687 MFLOPs	Param	6.04	5.18	**1.24**	1.79	3.30	1.63

Summary

- **No particular distribution of filters guarantees the best accuracy on all tasks.**
  Promising filter distributions for small datasets should be carefully extrapolated to big ones. In the opposite direction, filter designs working well on extensive datasets should be adapted to perform efficiently on different ones.
- **Exploring novel filter distributions has advantages that go beyond the domain of image classification.**
  Redistributing templates enhances performance and reduces resources for most of the models and domains tested.

References

[1] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.
2] Joseph Lin Chu and Adam Krzyżak. Analysis of feature maps selection in supervised learning using convolutional neural networks. In Canadian Conference on Artificial Intelligence, pages 59–70. Springer, 2014.
[3] Ihsan Ullah and Alfredo Petrosino. About pyramid structure in convolutional neural networks. In 2016 International joint conference on neural networks (IJCNN), pages 1318–1324. IEEE, 2016.
[4] Joseph Lin Chu and Adam Krzyżak. Analysis of feature maps selection in supervised learning using convolutional neural networks. In Canadian Conference on Artificial Intelligence, pages 59–70. Springer, 2014.
[5] Zhonghui You, Kun Yan, Jinmian Ye, Meng Ma, and Ping Wang. Gate decorator: Global filter pruning method for accelerating deep convolutional neural networks. In Advances in Neural Information Processing Systems, pages 2130–2141, 2019
[6] Ariel Gordon, Elad Eban, Ofir Nachum, Bo Chen, Hao Wu, Tien-Ju Yang, and Edward Choi. Morphnet:Fast & simple resource-constrained structure learning of deep networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1586–1595, 2018