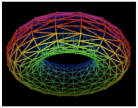
A Topological approach to convolutional neural networks

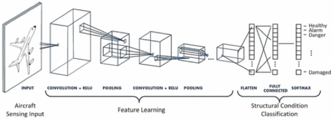
CycleCNN

Louis Maestrati

25/01/2021

CycleCNN: between Algebraic Topology and Deep Learning

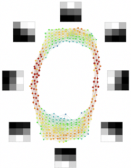




Algebraic Topology

CycleCNN

Deep Learning



Viewing CNN spatial filters as topological shape in $\mathbb{R}^9$

Homology in convolutional networks: Homology in CNNs

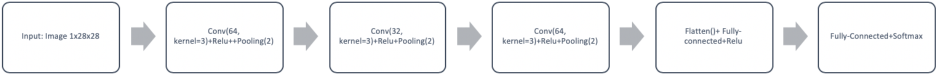


Figure: Convolutional neural network studied in [Carlson, 2018], translated from the formalism introduced in the paper.

Training on MNIST and CIFAR-10:

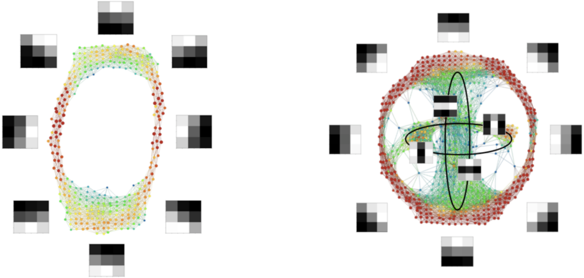


Figure: **1st Figure**: Mapper Model applied to the 1st layer of the network trained on MNIST , **2nd Figure**: Mapper Model applied to the 1st layer trained on CIFAR-10 (Figures from [Carlson, 2018])

Homology in convolutional networks: Homology in CNNs

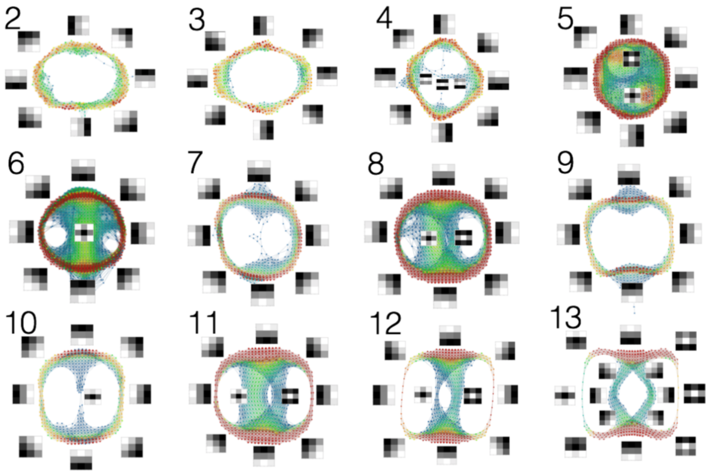


Figure: Mapper Model applied to the layers VGG-16 trained on ImageNet (Figures from de [Carlson, 2018])

Hypothesis and Motivations

Observations
The set of spatial filters tend to form ellipsoids in $\mathbb{R}^9$

Hypothesis
The set of kernels tend to form ellipsoids in $\mathbb{R}^{9 \times \# InputChannels}$

CycleCNN: Motivations

Observations

The set of spatial filters tend to form ellipsoids in $\mathbb{R}^9$

Hypothesis

The set of kernels tend to form circles in $\mathbb{R}^{9 \times \#InputChannels}$

We force these kernels to evolve from scratch on such ellipsoids

The hope is twice:

1. **Faster Training**: the optimal topological shape is already encoded in the architecture.
2. **Less parameters**: the gain in parameters need -and though in computational memory- is very important.

CycleCNN: Architecture

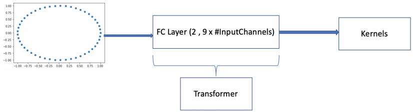


Figure: Diagram of the circle transformation in the kernels space $\mathbb{R}^9$

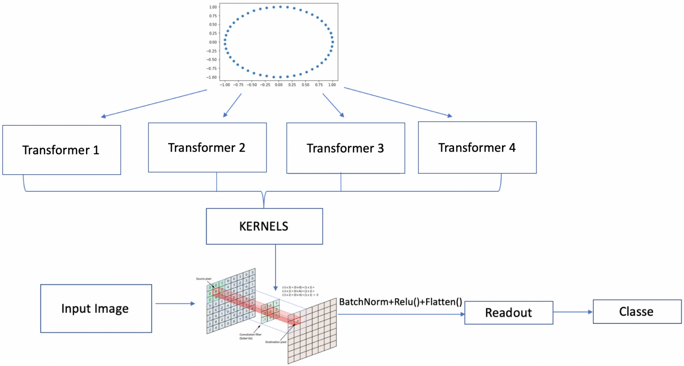


Figure: Diagram of a CycleCNN with one layer using 4 circles.

CycleCNN: Several layers model

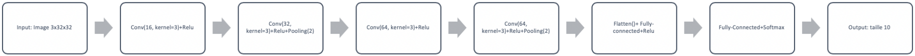


Figure: Diagram of the model used for comparisions.

Put in competition with avec:

- 1st Layer: 1 circle of kernels with 16 points (resulting in 16 kernels)
- 2nd Layer: 2 circles of kernels with 16 points (resulting in 32 kernels)
- 3rd cLayer: 4 circles of kernels with 16 points (resulting in 64 kernels)
- 4th layer: 2 cercles of kernels with 32 points (resulting in 64 kernels)

Our model uses $2 \times 3 \times 9 + 2 \times 2 \times 16 \times 9 + 2 \times 4 \times 32 \times 9 + 2 \times 2 \times 64 \times 9 = 5238$ parameters for convolution, whereas $16 \times 3 \times 9 + 32 \times 16 \times 9 + 64 \times 32 \times 9 + 64 \times 64 \times 9 = 60336$ parameters for the classic CNN.

Some results

#Classic kernels	#Cyclic kernels	Train	Test	#parameters CNN	transformer size	lr
64	0	85.2 %	75 %	18432	x	0.05
64	0	85.5 %	74.5 %	18432	x	0.01
48	16	**84.8** %	**74** %	13918	3	0.01/0.001
48	16	84.5 %	73.8 %	13851	1	0.01/0.001
48	0	84 %	73.2 %	13824	x	0.01
32	32	**82.3** %	**73.5** %	9404	3	0.01/0.001
32	32	81.5 %	72.8 %	9270	1	0.01/0.001
32	0	80.8 %	72.2 %	9216	x	0.01
16	48	75 %	70.6 %	4890	3	0.01/0.001
16	48	**79** %	**71.2** %	4689	1	0.01/0.001
16	0	73.5 %	69.3 %	4608	x	0.01
0	64	54 %	52 %	376	3	0.01/0.001
0	64	72 %	60 %	108	1	0.01/0.001

Figure: Results for a layer parameterized with different amounts Classic Kernels and Cyclic kernels. The model can be found in Annexe. The Cycle kernels are constructed with 4 circles.

Conclusion and future work

- Encouraging results, with few parameters we are able to increase accuracy but more a "boosting" method than a new CNN parameterization
- Topological circles are not necessarily ellipsoids: the notion of "thickness" is lost $\rightarrow$ future improvement
- initialization and learning rates for the transformers are critical choices

Références

Gunnar Carlsson, Rickard Bru el Gabrielsso.
Topological Approaches to Deep Learning.
arXiv:1811.01122, 2018.

Frédéric Chazal, Bertrand Michel.
An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists..
arXiv preprint arXiv:1710.04019v1, 2017.

Topaz CM, Ziegelmeier L, Halverson T.
Topological Data Analysis of Biological Aggregation Models..
PLoS ONE 10(5): e0126383., 2015.