Convolutional Neural Networks using separable filters

Petrescu Viviana
EPFL
viviana.petrescu@epfl.ch

Abstract

CNN are the new state-of-the-art machine learning techniques which achieved best results on various vision tasks ranging from the large scale ImageNet challenge to segmentation and to bio medical imaging. This technical report presents initial results on using CNN with separable filters for speeding up testing time.

1 Introduction

Although proven to be very powerful, CNN are much slower than their counter parts SVN or Random Forests. During one forward pass, the computational complexity of evaluation one image os size WxH with J filters of size k_1xk_2 is $O(WHJk_1k2)$. If the 2d filters are decomposed into separable 2d filters, the complexity becomes $O(WH(J + k_1 + k2)$. Thus, we obtain a speedup if $k << \frac{Jk_1k_2}{J+k_1+k_2}$.

Picture from

The theoretical rank of a 3D tensor R is less or equal with $prod$. This means that for 'tall' set of tensors like the ones present in CNNs $J \times k_1 \times k_2$ with $k_1 \times k_2 J$ we should expect the rank to be equal with the product of the two kernel dimensions. Some inequalities that hold for the theoretical rank are presented below: $\chi \in \mathbb{R}^{I \times J \times K}$. The typical rank is any rank that occurs with probability greater than 0. FFT Other common typical ranks:

2 Speeding up CNN

For the sake of completeness, some approaches that were previously tried were on [**?**], which prove speedup for CNNs for 3 different schemes, one of which it is similar with our approach but using a different optimization algorithm for obtaining the separable filters. Other approached would be to heavily employ the GPU and even employ the FGPA, which are harder to program [1] Talk about current available programming tools - Caffee increasingly popular - Theano symbolic based - Torch7 - mathLib

Talk about CNN with separable filters and with GPU, fgpa etc

3 Experiments

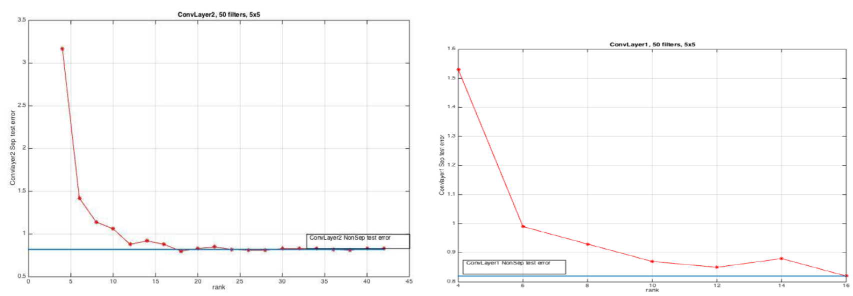
3.1 MNIST

MNIST consists of a curated set of postal digits images, of size 28x28. The training set contains 60000 samples, validation 10000 and testing 10000 samples. The approached of YannLecun[98] and Ciresan[] proved very good performance. The reference Theano model of Theano achieves 0.82 error rates.

Layer	Type	Maps and neurons	Kernel size
0	input	1 map of 28x28	
1	convolutional	20 maps of 24x24	5x5
2	max pooling	20 maps of 12x12	2x2
3	convolutional	50 maps of 8x8	5x5
4	max pooling	50 maps of 4x4	2x2
5	fully conntected	500	
6	fully conntected	2 neurons	

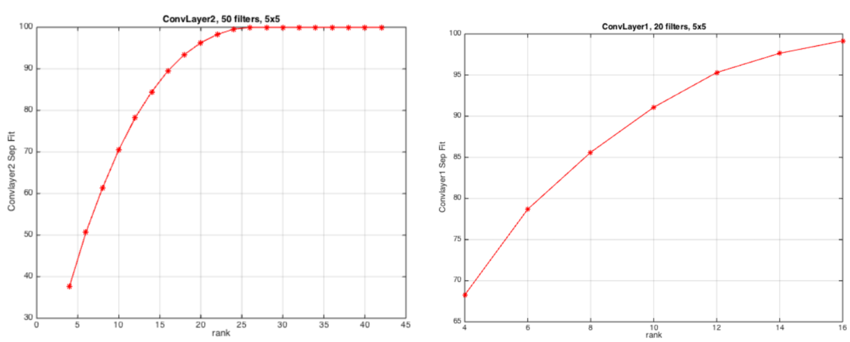
Table 1: Small CNN for MNIST set

3.1.1 Model 1



(a) (b)

Figure 1: Distribution of listening counts



(a) (b)

Figure 2: Distribution of listening counts

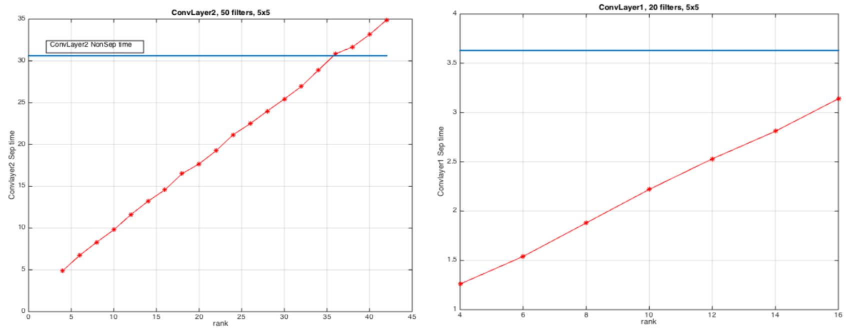
3.1.2 Model 2

3.2 Mitocondria

3.3 ImageNet

4 Conclusions

Theoretical bounds of the separabe filters choice are proven. Impreved recognition on mitocondria set.



(a) (b)

Figure 3: Distribution of listening counts

Layer	Type	Maps and neurons	Kernel size
0	input	1 map of 28x28	
1	convolutional	20 maps of 24x24	9x9
2	max pooling	20 maps of 12x12	2x2
3	convolutional	50 maps of 8x8	9x9
4	max pooling	50 maps of 4x4	2x2
5	fully conntected	500	
6	fully conntected	2 neurons	

Table 2: Bigger CNN for MNIST set

Acknowledgments

We would like to thank Simon Arosi that helped me a lot during this project giving me constant feedback.

References

[1] Yann LeCun, Koray Kavukcuoglu, and Clément Farabet. Convolutional networks and applications in vision. In *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pages 253–256. IEEE, 2010.

[2] Yunhong Zhou, Dennis Wilkinson, Robert Schreiber, and Rong Pan. Large-scale parallel collaborative filtering for the netflix prize. In *Proceedings of the 4th International Conference on Algorithmic Aspects in Information and Management*, AAIM '08, pages 337–348, 2008.