GaborNet

**Abstract**

This research on deep convolutional neural networks proposes a modified architecture that focuses on improving convergence and reducing training complexity. The filters in the first layer of network are constrained to fit the Gabor function. The parameters of Gabor functions are learnable and updated by standard backpropagation techniques. The proposed architecture was tested on several datasets and outperformed the common convolutional networks.

**Introduction**

Convolutional neural networks (CNNs) have various applications, especially in computer vision. They have attracted significant attention due to its ability to be trained end-to-end and capability of learning outstanding feature representations from raw image data. In 2015 CNN showed human-like performance on ImageNet dataset [1]. Unlike classic computer vision methods, neural network is a data driven algorithm, that learns robust representations from data, but usually at the cost of training excessive number of parameters (or weights). Additionally, convergence of neural network depends on parameters initialization. Usually the weights are initialized by uniform or normal distribution. But it causes convergence problem and makes difficult to train very deep CNN.

Glorot and Bengio (2010) [2] proposed a formula, that estimates standard deviation for a CNN layer based on layer input and output size. In [1] new method of initialization, known as He initialization, was introduced, which also suggests using normal or uniform distribution with parameters based on neural network topology and type of activation functions. However, these approaches focused mainly on dealing with training deep CNN. Hence additional constraints and initialization methods should be applied to convolutional neural networks’ weights to overcome slow convergence speed. Specifically, this could be achieved by generating weights of neural network by family of parameterized filters.

Gabor filters [3] are widely used in computer vision. They are based on sinusoidal plane wave with particular frequency and orientation, which allows them to extract spatial frequency structures from images [4]. Additionally, these filters proved to be appropriate for texture representation and face detection tasks [5]. A few works have explored Gabor filters for CNN. In [6] Gabor filters were used as preprocessing tool to generate Gabor features then using it as an input to a CNN, in [7] first or second layer of CNN was set as a constant Gabor filter bank, thus reducing number of trainable parameters of the network, and in [8] Convolutional Gabor orientation Filters were introduced, a special structure that modulates convolutional layers with learnable parameters by non-learnable Gabor filter bank. However, they do not integrate Gabor filters into backpropagation algorithm.

In this paper we propose using Gabor Layer as a first layer in deep convolutional network. Gabor Layer is a convolutional layer, which filters are constrained to fit Gabor functions (Gabor filters). The parameters of Gabor functions are initialized from filter bank, proposed in [9], and during training process the filter parameters are updated by standard backpropagation algorithm. This approach aims to increase robustness of learned feature representations and to reduce training complexity of neural networks. Gabor Layer is implemented on basic elements of CNNs and can be easily integrated into any deep CNN architecture.

**Theory**

**CNN**

Convolutional neural networks usually consist of convolutional layers and fully connected layers, with nonlinear activation function applied at the end of each layer. Pooling and dropout layers may be included in the architecture to avoid overfitting. Convolutional layers reduce memory usage and increase performance by using the same filters for each pixel of the image. The standard architecture of CNN is shown in Fig.1.

The training process of CNN usually is carried out by backpropagation algorithm [11: Goodfellow, Ian; Bengio, Yoshua; Courville, Aaaron (2016) Deep Learning. MIT Press. p. 196. [ISBN](https://en.wikipedia.org/wiki/International_Standard_Book_Number) [9780262035613](https://en.wikipedia.org/wiki/Special:BookSources/9780262035613)] with active dropout layers. The evaluation process is much simpler, because it requires only forward propagation, dropout layers is inactive during evaluation process.

**Gabor function**

Gabor function is a complex sinusoid modulated by Gaussian i.e.

However, in this work we used real values of Gabor function i.e.

Gabor filters proved to be an efficient tool for extracting spatially localized spectral features, which are used in various pattern analysis applications. In [10] was shown that deep CNN trained on real-life images tends to learn first and second convolutional layers contain mostly Gabor-like filters. Filters of first and second layers of AlexNet is shown in Fig.2. This corroborates an idea of using Gabor filters in first layer of CNN.

To initialize Gabor filters in Gabor Layer, approach proposed in [11] was used. Frequencies and orientations of the Gabor filters are obtained by the following equations:

The is set by , which allows to define relationship between and . The is set by uniform distribution U(0, ). During the training process these parameters are learned by backpropagation algorithms.

Gabor function formula -

derivative of Gabor function -

**Experiment**

* table with nets architecture (size of Gabor layers)
* table with info about datasets
* accuracy/epoch plots

weights initialization, datasets, networks architectures, plots, accuracy

References

<https://www.researchgate.net/publication/272912407_Face_recognition_using_Gabor_filter_bank_kernel_principal_component_analysis_and_support_vector_machine>

<http://mohammadmahoor.com/affectnet/>

<https://www.microsoft.com/en-us/research/publication/asirra-a-captcha-that-exploits-interest-aligned-manual-image-categorization/> , <https://www.kaggle.com/c/dogs-vs-cats>

<http://www.image-net.org/>

<http://yann.lecun.com/exdb/publis/pdf/lecun-89e.pdf>

<https://www.matlabi.ir/wp-content/uploads/bank_papers/g_paper/g105_Matlabi.ir_A%20New%20Facial%20Expression%20Recognition%20Method%20Based%20on%20Local%20Gabor%20Filter%20Bank%20and%20PCA%20plus%20LDA.pdf>

# References

|  |  |
| --- | --- |
| [1] | K. He, X. Zhang, S. Ren and J. Sun, "Delving Deep into Rectifiers:," 6 Feb 2015. [Online]. Available: https://arxiv.org/pdf/1502.01852.pdf. |
| [2] | X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010. |
| [3] | D. Gabor, "Theory of Communication," *Journal of the Institution of Electrical Engineers,* pp. 429-441, 1946. |
| [4] | A. K. Jain, N. K. Ratha and S. Lakshmanan, "Object detection using gabor filters," *Pattern Recognition,* vol. 30, no. 2, pp. 295-309, 1997. |
| [5] | Huang, L.-L. &. Shimizu, A. &. Kobatake and Hidefumi, "Classification-Based Face Detection Using Gabor Filter Features," in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 2014. |
| [6] | K. Bogdan, "Face Detection Using Convolutional Neural Networks and Gabor Filters," 0302-9743, 2005. |
| [7] | A. Calderón, S. R. Ovalle and J. Victorino, "Handwritten Digit Recognition using Convolutional Neural Networks and Gabor filters," 2003. |
| [8] | S. Luan, B. Zhang and S. Zhou, "Gabor Convolutional Networks," 29 Jan 2018. [Online]. Available: https://arxiv.org/pdf/1705.01450.pdf. |
| [9] | H.-B. Deng, L.-W. Jin, L.-X. Zhen and J.-C. Huang, "A New Facial Expression Recognition Method Based on," *International Journal of Information Technology,* vol. 11, no. 11, pp. 86-96, 2005. |
| [10] | A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *NIPS2012*, 2012. |
| [11] | S. Meshgini, A. Aghagolzadeh and H. Seyedarabi, "Face recognition using Gabor filter bank, kernel principal component analysis and support vector machine," *International Journal of Computer Theory and Engineering,* vol. 4, pp. 767-771, 2012. |