zbMATH Open

THE FIRST RESOURCE FOR MATHEMATICS

Tygert, Mark; **Bruna, Joan**; **Chintala, Soumith**; **LeCun, Yann**; **Piantino, Serkan**; **Szlam, Arthur**

A mathematical motivation for complex-valued convolutional networks. (English)

Zbl 07062543

Neural Comput. 28, No. 5, 815-825 (2016).

Summary: A complex-valued convolutional network (convnet) implements the repeated application of the following composition of three operations, recursively applying the composition to an input vector of nonnegative real numbers: (1) convolution with complex-valued vectors, followed by (2) taking the absolute value of every entry of the resulting vectors, followed by (3) local averaging. For processing real-valued random vectors, complex-valued convnets can be viewed as data-driven multiscale windowed power spectra, data-driven multiscale windowed absolute spectra, data-driven multiwavelet absolute values, or (in their most general configuration) data-driven nonlinear multiwavelet packets. Indeed, complex-valued convnets can calculate multiscale windowed spectra when the convnet filters are windowed complex-valued exponentials. Standard real-valued convnets, using rectified linear units (ReLUs), sigmoidal (e.g., logistic or tanh) nonlinearities, or max pooling, for example, do not obviously exhibit the same exact correspondence with data-driven wavelets (whereas for complex-valued convnets, the correspondence is much more than just a vague analogy). Courtesy of the exact correspondence, the remarkably rich and rigorous body of mathematical analysis for wavelets applies directly to (complex-valued) convnets.

MSC:

68 Computer science

Software:

CIFAR; SIFT; Steerable pyramid; SURF

Full Text: DOI

References:

[1] Bay, H., Ess, A., Tuytelaars, T., \& Gool, L. V. (2008). Speeded-up robust features (SURF). Computer Vision Image Understanding, 110(3), 346-359. ,

[2] Bruna, J., \& Mallat, S. (2013). Invariant scattering convolutional networks. IEEE Trans. Pattern Analysis Machine Intel., 35(8), 1872-1886. ,

[3] Bruna, J., Mallat, S., Bacry, E., \& Muzy, J.-F. (2015). Intermittent process analysis with scattering moments. Ann. Statist., 43(1), 323-351. , · Zbl 1308.62168

[4] Chintala, S., Ranzato, M., Szlam, A., Tian, Y., Tygert, M., \& Zaremba, W. (2015). Scale-invariant learning and convolutional networks (Tech. Rep.). 1506.08230, arXiv. · Zbl 1392.68340

[5] Coifman, R. R., \& Donoho, D. (1995). Translation-invariant denoising. In A. Antoniadis \& G. Oppenheim (Eds.), Wavelets and statistics (pp. 125-150). New York: Springer. , · Zbl 0866.94008

[6] Coifman, R. R., Meyer, Y., Quake, S., \& Wickerhauser, M. V. (1994). Signal processing and compression with wavelet packets. In J. S. Byrnes, J. L. Byrnes, K. A. Hargreaves, \& K. Berry (Eds.), Wavelets and their applications (pp. 363-379). New York: Springer. , · Zbl 0818.94005

[7] Dalal, N., \& Triggs, B. (2005). Histograms of oriented gradients for human detection. In Proceedings of the IEEE Computer Society Conf. Computer Vision and Pattern Recognition 2005 (vol. 1, pp. 886-893). Piscataway, NJ: IEEE. ,

[8] Daubechies, I. (1992). Ten lectures on wavelets. Philadelphia: SIAM. , · Zbl 0776.42018

[9] Donoho, D., Mallat, S., von Sachs, R., \& Samuelides, Y. (2003). Locally stationary covariance and signal estimation with macrotiles. IEEE Trans. Signal Processing, 51(3), 614-627. , · Zbl 1369.94132

[10] Haensch, R., \& Hellwich, O. (2010). Complex-valued convolutional neural networks for object detection in PolSAR data. In Proceedings of the 8th European Conf. EUSAR (pp. 1-4). Piscataway, NJ: IEEE.

[11] Krizhevsky, A. (2009). Learning multiple layers of features from tiny images. Master's thesis, University of Toronto.

[12] Krizhevsky, A., Sutskever, I., \& Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, \& K. Q. Weinberger (Eds.), Advances in neural information processing systems, 25 (pp. 1097-1105). Red Hook, NY: Curran.

[13] LeCun, Y., Bengio, Y., \& Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444. ,

[14] LeCun, Y., Bottou, L., Bengio, Y., \& Haffner, P. (1998). Gradient-based learning applied to document recognition. Proc. IEEE, 86(11), 2278-2324. ,

[15] Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Proceedings of the 7th IEEE Internat. Conf. Computer Vision (vol. 2, pp. 1150-1157). Piscataway, N.J: IEEE.

[16] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. Internat. J. Computer Vision, 60(2), 91-110. ,

[17] Mallat, S. (2008). A wavelet tour of signal processing: The sparse way (3rd ed.). Orlando, FL: Academic Press. · Zbl 1170.94003

[18] Mallat, S. (2010). Recursive interferometric representations. In Proc. of the EUSIPCO Conf. 2010 (pp. 716-720). Piscataway, NJ: IEEE.

[19] Mehta, P., \& Schwab, D. J. (2014). An exact mapping between the variational renormalization group and deep learning. (Tech. Rep.). 1410.3831, arXiv.

[20] Meyer, Y. (1993). Wavelets and operators. Cambridge: Cambridge University Press. , · Zbl 0810.42015

[21] Meyer, Y., \& Coifman, R. R. (1997). Wavelets: Calderón-Zygmund and multilinear operators. Cambridge: Cambridge University Press. · Zbl 0916.42023

[22] Oyallon, E., \& Mallat, S. (2015). Deep roto-translation scattering for object classification. In Proceedings of the IEEE Computer Society Conf. Computer Vision and Pattern Recognition 2015 (vol. 1, pp. 2865-2873). Piscataway, NJ: IEEE. ,

[23] Poggio, T., Mutch, J., Leibo, J., Rosasco, L., \& Tacchetti, A. (2012). The computational magic of the ventral stream: Sketch of a theory (and why some deep architectures work) (Tech. Rep. MIT-CSAIL-TR-2012-035), Cambridge, MA: MIT CSAIL.

[24] Rabiner, L. R., \& Schafer, R. W. (2007). Introduction to digital speech processing.Hanover, MA: Now Publishers. · Zbl 1162.94003

[25] Saito, N.& Coifman, R. R. (1995). Local discriminant bases and their applications. J. Math. Imaging Vision, 5(4), 337-358. , · Zbl 0863.94004

[26] Simoncelli, E. P., \& Freeman, W. T. (1995). The steerable pyramid: A flexible architecture for multi-scale derivative computation. In Proceedings of the Internat. Conf. Image Processing 1995 (vol. 3, pp. 444-447). Piscataway, NJ: IEEE.

[27] Srivastava, A., Lee, A. B., Simoncelli, E. P., \& Zhu, S. (2003). On advances in statistical modeling of natural images. J. Math. Imaging Vision, 18(1), 17-33. , · Zbl 1033.68133