AlexNet

AlexNet is the name of a convolutional neural network, designed by Alex Krizhevsky,[1] and published with Ilya Sutskever and Krizhevsky's PhD advisor Geoffrey Hinton,[2][3] who was originally resistant to the idea of his student.[1][4]

AlexNet competed in the ImageNet Large Scale Visual Recognition Challenge[5] on September 30, 2012. The network achieved a top-5 error of 15.3%, more than 10.8 percentage points lower than that of the runner up. The original paper's primary result was that the depth of the model was essential for its high performance, which was computationally expensive, but made feasible due to the utilization of graphics processing units (GPUs) during training.[5]

Contents

Historic context

AlexNet was not the first fast GPU-implementation of a CNN to win an image recognition contest. A CNN on GPU by K. Chellapilla et al. (2006) was 4 times faster than an equivalent implementation on CPU.[6] A deep CNN of Dan Ciresan et al. (2011) at IDSIA was already 60 times faster[7] and achieved superhuman performance in August 2011.[8] Between May 15, 2011 and September 10, 2012, their CNN won no fewer than four image competitions.[9][10] They also significantly improved on the best performance in the literature for multiple image databases.[11]

According to the AlexNet paper,[5] Ciresan's earlier net is "somewhat similar." Both were originally written with CUDA to run with GPU support. In fact, both are actually just variants of the CNN designs introduced by Yann LeCun et al. (1989)[12][13] who applied the backpropagation algorithm to a variant of Kunihiko Fukushima's original CNN architecture called "neocognitron."[14][15] The architecture was later modified by J. Weng's method called max-pooling.[16][10]

In 2015, AlexNet was outperformed by Microsoft's very deep CNN with over 100 layers, which won the ImageNet 2015 contest.[17]

Network design

AlexNet contained eight layers; the first five were convolutional layers, some of them followed by max-pooling layers, and the last three were fully connected layers.[5] It used the non-saturating ReLU activation function, which showed improved training performance over tanh and sigmoid.[5]

Influence

AlexNet is considered one of the most influential papers published in computer vision, having spurred many more papers published employing CNNs and GPUs to accelerate deep learning.[18] As of 2019, the AlexNet paper has been cited over 47,000 times.

Alex Krizhevsky

Alex Krizhevsky (born in Ukraine, raised in Canada) is a computer scientist most noted for his work on artificial neural networks and deep learning. Shortly after having won the ImageNet challenge 2012 through AlexNet, he and his colleagues sold their startup DNN Research Inc. to Google.[1] Krizhevsky left Google in September 2017 when he lost interest in the work.[1] At the company Dessa, Krizhevsky will advise and help research new deep-learning techniques.[1] Many of his numerous papers on machine learning and computer vision are frequently cited by other researchers.[19]

References

1. Dave Gershgorn (18 June 2018). "The inside story of how AI got good enough to dominate Silicon Valley" (https://qz.com/1307091/the-inside-story-of-how-ai-got-good-enough-to-dominate-silicon-valley/). *Quartz*. Retrieved 5 October 2018.
2. "The data that transformed AI research—and possibly the world" (https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/).
3. "ILSVRC2012 Results" (http://www.image-net.org/challenges/LSVRC/2012/results.html).
4. Krizhevsky, Alex. "ImageNet Classification with Deep Convolutional Neural Networks" (http://www.image-net.org/challenges/LSVRC/2012/supervision.pdf) (PDF). Retrieved 17 November 2013.
5. Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey E. (2017-05-24). "ImageNet classification with deep convolutional neural networks" (https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf) (PDF). *Communications of the ACM*. **60** (6): 84–90. doi:10.1145/3065386 (https://doi.org/10.1145%2F3065386). ISSN 0001-0782 (https://www.worldcat.org/issn/0001-0782).
6. Kumar Chellapilla; Sid Puri; Patrice Simard (2006). "High Performance Convolutional Neural Networks for Document Processing" (https://hal.inria.fr/inria-00112631/document). In Lorette, Guy (ed.). *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft.
7. Ciresan, Dan; Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber (2011). "Flexible, High Performance Convolutional Neural Networks for Image Classification" (http://www.idsia.ch/~juergen/ijcai2011.pdf) (PDF). *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence-Volume Volume Two*. **2**: 1237–1242. Retrieved 17 November 2013.
8. "IJCNN 2011 Competition result table" (http://benchmark.ini.rub.de/?section=gtsrb&subsection=results). *OFFICIAL IJCNN2011 COMPETITION*. 2010. Retrieved 2019-01-14.
9. Schmidhuber, Jürgen (17 March 2017). "History of computer vision contests won by deep CNNs on GPU" (http://people.idsia.ch/~juergen/computer-vision-contests-won-by-gpu-cnns.html). Retrieved 14 January 2019.
10. Schmidhuber, Jürgen (2015). "Deep Learning" (http://www.scholarpedia.org/article/Deep_Learning). *Scholarpedia*. **10** (11): 1527–54. CiteSeerX 10.1.1.76.1541 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.76.1541). doi:10.1162/neco.2006.18.7.1527 (https://doi.org/10.1162%2Fneco.2006.18.7.1527). PMID 16764513 (https://www.ncbi.nlm.nih.gov/pubmed/16764513).
11. Ciresan, Dan; Meier, Ueli; Schmidhuber, Jürgen (June 2012). *Multi-column deep neural networks for image classification*. *2012 IEEE Conference on Computer Vision and Pattern Recognition*. New York, NY: Institute of Electrical and Electronics Engineers (IEEE). pp. 3642–3649. arXiv:1202.2745 (https://arxiv.org/abs/1202.2745). CiteSeerX 10.1.1.300.3283 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.300.3283). doi:10.1109/CVPR.2012.6248110 (https://doi.org/10.1109%2FCVPR.2012.6248110). ISBN 978-1-4673-1226-4. OCLC 812295155 (https://www.worldcat.org/oclc/812295155).
12. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation Applied to Handwritten Zip Code Recognition (http://yann.lecun.com/exdb/publis/pdf/lecun-89e.pdf); AT&T Bell Laboratories

13. LeCun, Yann; Léon Bottou; Yoshua Bengio; Patrick Haffner (1998). "Gradient-based learning applied to document recognition" (http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf) (PDF). *Proceedings of the IEEE*. **86** (11): 2278–2324. CiteSeerX 10.1.1.32.9552 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.9552). doi:10.1109/5.726791 (https://doi.org/10.1109%2F5.726791). Retrieved October 7, 2016.
14. Fukushima, K. (2007). "Neocognitron". *Scholarpedia*. **2** (1): 1717. doi:10.4249/scholarpedia.1717 (https://doi.org/10.4249%2Fscholarpedia.1717).
15. Fukushima, Kunihiko (1980). "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position" (http://www.cs.princeton.edu/courses/archive/spr08/cos598B/Readings/Fukushima1980.pdf) (PDF). *Biological Cybernetics*. **36** (4): 193–202. doi:10.1007/BF00344251 (https://doi.org/10.1007%2FBF00344251). PMID 7370364 (https://www.ncbi.nlm.nih.gov/pubmed/7370364). Retrieved 16 November 2013.
16. Weng, J; Ahuja, N; Huang, TS (1993). "Learning recognition and segmentation of 3-D objects from 2-D images". *Proc. 4th International Conf. Computer Vision*: 121–128.
17. He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian (2016). "Deep Residual Learning for Image Recognition". *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
18. Deshpande, Adit. "The 9 Deep Learning Papers You Need To Know About (Understanding CNNs Part 3)" (https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html). *adeshpande3.github.io*. Retrieved 2018-12-04.
19. "Alex Krizhevsky" (https://scholar.google.com/citations?user=xegzhJcAAAAJ). *Google Scholar Citations*.
