Deep Learning with TensorFlow
http://cvml.ist.ac.at/courses/DLWT_W18

Lecture 6:
AlexNet
AlexNet with Vyacheslav Li

December 17, 2018
Outline

• Historical introduction
• Architecture of AlexNet
• Key characteristics of AlexNet
• After AlexNet breakthrough
Outline

• Historical introduction
• Architecture of AlexNet
• Key characteristics of AlexNet
• After AlexNet breakthrough
ImageNet Classification with Deep Convolutional Neural Networks (2012)

Alex Krizhevsky              Ilya Sutskever               Geoffry Hinton

According to Google Scholar it has 32,680 citations
~ 15 million labeled high resolution images
~ 22,000 categories
ILSVRC

- 1000 images in each of 1000 categories
- **1.2 million** training images
- 50,000 validation images
- 150,000 testing images
Revolution of Depth

ImageNet Image Classification Top-5 Error(%)
1943: McCulloch & Pitts show that neurons can be combined to construct a Turing machine (using ANDs, ORs, & NOTs).
1985: The backpropagation algorithm by Geoffrey Hinton et al revitalizes the field
1998: CNNs with Backpropagation for document analysis by Yan LeCun
2012: AlexNet by Alex Krizhevsky in 2012
• Historical introduction
• **Architecture of AlexNet**
• Key characteristics of AlexNet
• After AlexNet breakthrough
62.3 million parameters and needs 1.1 billion computation units in a forward pass
Architecture of AlexNet
Outline

• Historical introduction
• Architecture of AlexNet
• **Key characteristics of AlexNet**
• After AlexNet breakthrough
ReLU

The ReLU activation was first used in AlexNet

\[ y = \max(0, x) \]
Trained on CIFAR – 10 dataset
Why ReLU faster?
Response normalization reduced top-1 and top-5 error rates by 1.4% and 1.2%, respectively.
Max pooling

\[
\begin{array}{cccc}
12 & 20 & 30 & 0 \\
8 & 12 & 2 & 0 \\
34 & 70 & 37 & 4 \\
112 & 100 & 25 & 12 \\
\end{array}
\]

\[2 \times 2 \text{ Max-Pool}\]

\[
\begin{array}{cc}
20 & 30 \\
112 & 37 \\
\end{array}
\]
Overlapping Max pooling

$S < Z$ instead of traditional $S=Z$ ($S$: stride, $Z$: size of filter matrix)

This scheme reduces the top-1 and top-5 error rates by 0.4% and 0.3%, respectively.
Reducing Overfitting

Data Augmentation
Reducing overfitting

Data Augmentation

Random Crops

256 → 227
Reducing overfitting

Data Augmentation

Increased the number of training data by factor of 2048!
Reducing overfitting

Data Augmentation

Original Image

Fancy PCA

Random Augmentation
Reducing overfitting

Data Augmentation

Principal Components Analysis was performed on the set of RGB pixel values. Varying intensities helped to reduce error by over 1%
Reducing overfitting

Drop out 0.5 of neurons during training
Used GPU to increase training speed

Double GPU reduced top-1 and top-5 error rates by **1.7%** and **1.2%**, respectively.
• Historical introduction
• Architecture of AlexNet
• Key characteristics of AlexNet
• After AlexNet breakthrough
Growing Use of Deep Learning at Google

Number of directories containing model description files

Across many products/areas
- Apps
- Maps
- Photos
- Gmail
- Speech
- Android
- YouTube
- Translation
- Robotics Research
- Image Understanding
- Natural Language Understanding
- Drug Discovery
Results

The graph shows the number of scientific papers over time for the keywords "Deep Learning" and "Deep Learning + Life Sciences" from 2005 to 2017. The trend for "Deep Learning + Life Sciences" starts from a negligible number in 2005 and sees a sharp increase after 2013, indicating a significant growth in research interest in this area.
References

- [https://medium.com/@smallfishbigsea/a-walk-through-of-alexnet-6cbd137a5637](https://medium.com/@smallfishbigsea/a-walk-through-of-alexnet-6cbd137a5637)
- [https://www.learnopencv.com/understanding-alexnet/](https://www.learnopencv.com/understanding-alexnet/)