CIFAR-10 Image Recognition

EE4305 Introduction to Fuzzy/Neural Systems

Mario Gini
Thomas Hayden

Contents

1 Introduction

T. HAYDEN

The CIFAR-10 dataset contains 60000 images bla bla.
Objectives of this project are: bla bla
Structure of the report is as follows: bla bla

2 Literature Review on Artificial Neural Networks

M. GINI

This section gives a literature review on the broad topic of artificial neural networks (ANN). A more specific review on ANN designed to classify the CIFAR-10 dataset is found in Section 3. The significance and applications of ANN will be reviewed in Section 2.1 while recent trends and accomplishments are discussed in Section 2.2.

2.1 Significance and Applications of Artificial Neural Networks

This subsection will illustrate the significance and applications of ANN. Increasing computer power shifted the focus of research towards deep ANN and similar architectures which are coined under the term ”deep learning”. These powerful deep ANN are used in a variety of applications[1][2].

ANN are significant because they can work as a black box model. For example, an ANN can be trained to classify images into different categories. The performance can be enhanced by data preprocessing, augmentation and mainly by finding an appropriate network architecture. No a-priori knowledge of the classification process itself is required. This makes deep ANN suited for applications where such knowledge is difficult to obtain. Character and speech recognition are two further problems which are very hard to . In speech recognition, deep ANN have been shown to outperform other methods on a variety of speech recognition benchmarks, sometimes by a large margin[3].

However, the fact that ANN do not incorporate much of a-priori knowledge is not only positive. As a consequence, a trained model gives little insight into its inner workings and optimal network architectures are basically found through a trial-and-error process. Most design guidelines for deep learning methods are therefore based on empirical knowledge instead of theoretical foundations.

There is effort going on to better understand the computations deep ANN perform at each laver, with a few interesting visualizations available[4][5]. In general, each layer extracts higher level features of the input data until the last layer finally classifies the input.

2.2 Recent Trends and Accomplishments

One of the first big accomplishments of NN, specifically of convolutional neural networks (CNN), was during the ImageNet competition 2012 when a deep network won[6].

Recent trends and accomplishments of neural networks are discussed in this subsection. Since neural networks are such a broad topics, only two recent accomplishments are looked at in detail: The AlphaGo computer program and adversial examples. AlphaGo is a great example to illustrate recent trends and is also considered one of the biggest accomplishments of deep neural networks up to date. Adversial examples can easily fool very different kinds of neural networks and are therefore a great example to show that enthusiasm maybe should be insulated.

AlphaGo is a computer program developed by DeepMind, a company owned by Google. It uses deep learning and is able to beat the world champion. This gained considerable media coverage since it is a feat experts thought would still be a decade away [CITE]. The game Go is considered the most complex board game, with a number of legal positions of in the order of magnitude of 10^{170}[7].

Dedicated hardware is developed to accelerate the training process of a deep ANN. The most notable example is the Tensor Processing Unit developed by Google which achieves a 15- to 30 fold performance compared to a contemporary GPU or CPU[8].

[9][10]

Interesting development: Adversial examples

Adversial examples are a recalcitrant problem. It is a potential security problem. Studying them can im[11][12]

3 Literature Review on the CIFAR-10 dataset

T. HAYDEN

The CIFAR-10 data set[13] is a well established data set in the machine learning community. It is challenging because it is a relatively small data set. Even so, excellent results, even exceeding human performance, have been obtained using a variety of CNN architectures[1].

[1]http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html

At the time of writing, the highest published result on the CIFAR-10 data set was achieved in 2015 using a CNN with a fractional max-pooling set up to allow for a deeper network[14]. Using a max-pooling architecture, an accuracy of 96.53% was obtained. This is considerably better than human performance which has an accuracy of around 94%[15].

3.1 Data Augmentation

Like many other machine learning problems image, image classification will almost always benefit from additional data[16]. However, even when restricted to a particular dataset such as CIFAR-10 it is possible to generate more data using a technique called data augmentation[17]. Data augmentation manipulates existing images to create 'new' data for use in training.

Common methods to augment images for use in machine learning include mirroring, rotation and image translation[6]. Using these techniques it is possible to train on a data set that can be several times larger than the original data set. The leading architectures all made use of image augmentation[14][18][19].

3.2 Fractional Max-pooling

3.3 Layer-sequential unit-variance (LSUV) initialization

3.4 The All Convolutional Net

4 MLP Classifier

M. GINI

This section should contain:

a) on the selection of the inputs and outputs of the MLP b) on the size of the training data c)on the training of the MLP d) on the performance of the MLP with different objective functions and optimization methods e)any other interesting observation that you think are pertinent (e.g. effect of learning rate on convergence speed).

4.1 Data Preprocessing and Augmentation

- Normalization

  The input data is normalized to lie within the range [0,1].

- Mean subtraction

  To further normalize the data, a the mean is subtracted on a per-pixel basis.

- Data augmentation

  Experience shows that a larger training data set increases network performance

4.2 Network Structure

- Basic structure

  Since this is a classification problem, parts of the network structure are fixed. The last layer consists of 10 nodes and a in a ”softmax” configuration. PICTURE of basic structure.

- Number of hidden layers/nodes
  Parameter search over 1-3 hidden layers, 1-500 neurons

4.3 Optimization of Further Network Parameters

- Different learning rates
- Different optimization methods

5 CNN network

M. GINI

6 Conclusion

M. GINI

Long story short: we completely aced our project BOOM

Bibliography

[1] Haohan Wang and Bhiksha Raj. A survey: Time travel in deep learning space: An introduction to deep learning models and how deep learning models evolved from the initial ideas. *arXiv preprint arXiv:1510.04781*, 2015.

[2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553): 436–444, 2015.

[3] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[4] Alexander Mordvintsev, Christopher Olah, and Mike Tyka. Inceptionism: Going Deeper into Neural Networks, June 2015. URL `http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html`.

[5] Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs, and Hod Lipson. Understanding neural networks through deep visualization. *arXiv preprint arXiv:1506.06579*, 2015.

[6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[7] John Tromp and Gunnar Farnebäck. Combinatorics of go. In *International Conference on Computers and Games*, pages 84–99. Springer, 2006.

[8] Norman P Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, et al. In-datacenter performance analysis of a tensor processing unit. *arXiv preprint arXiv:1704.04760*, 2017.

[9] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[10] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[11] Anh Nguyen, Jason Yosinski, and Jeff Clune. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[12] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.

[13] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. 2009.

[14] Benjamin Graham. Fractional max-pooling. *arXiv preprint arXiv:1412.6071*, 2014.

[15] Andrej Karpathy. Lessons learned from manually classifying cifar-10. *Published online at http://karpathy. github. io/2011/04/27/manually-classifying-cifar10*, 2011.

[16] Alon Halevy, Peter Norvig, and Fernando Pereira. The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2):8–12, 2009.

[17] Xiaodong Cui, Vaibhava Goel, and Brian Kingsbury. Data augmentation for deep neural network acoustic modeling. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 23(9):1469–1477, 2015.

[18] Dmytro Mishkin and Jiri Matas. All you need is a good init. *arXiv preprint arXiv:1511.06422*, 2015.

[19] Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*, 2014.