Deep Learning Methods
Project 2

Jakub Kała & Krzysztof Spaliński

April 2020

Abstract. The scope of the second project for Deep Learning Methods course are convolutional neural networks used in image recognition tasks. The goal is to implement and explore several network architectures and data augmentation methods. Model assessment will be carried out based on CIFAR-10 dataset.

1 Documentation

TODO

2 CIFAR-10 Dataset

CIFAR-10 [KNH] is a computer-vision dataset used for object recognition. In our case, it is a subset of the 80 million tiny images dataset and consists of 60,000 32x32 color images containing one of 10 object classes, with 6000 images per class. It was collected by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton.

Dataset has been split to training and test set (50k and 10k images respectively). In order to discourage certain forms of cheating (such as hand labeling) 290,000 junk images have been added to the test set. These images are ignored in the scoring.The classes are completely mutually exclusive. The label classes in the dataset are:

- airplane
- automobile
- bird
- cat
- deer
- dog
- frog
- horse
- ship
- truck

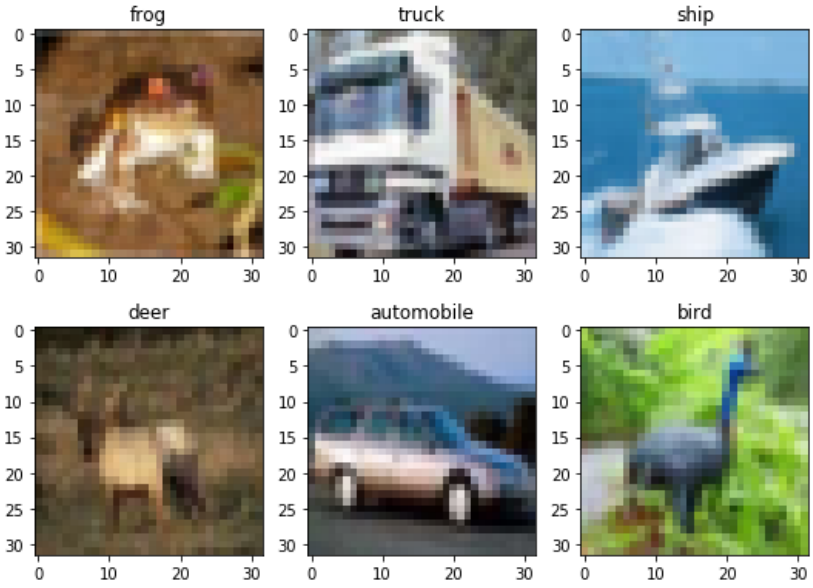

Figure 1: Example images from CIFAR-10 dataset.

3 Convolutional Neural Networks

Covolutional Neural Networks, CNNs, are neural networks employing convolution operation. They are really specialized for data with grid-like topology [GBC].

Convolution is terms of neural network is operation on two tensors: input and kernel (sometimes called a filter). Researchers developed many kernels that can be used for linear image processing, for example for blurring, sharpening, detecting edges etc. In deep learning, elements of filters are trainable values.

4 Basic CNN architecures

proste architektury, rozniace sie tylko liczba warstw i zdolnosci uczenia sie tych sieci?

5 Neural network with residual blocks

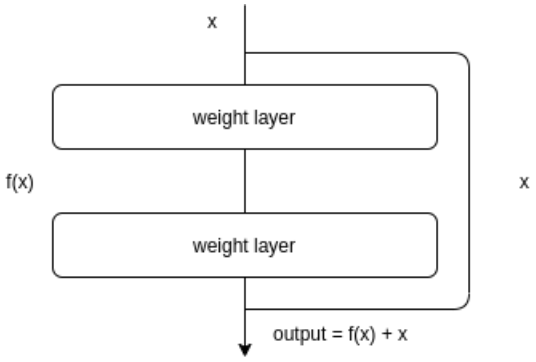

Figure 2: Residual block example with 2 skipped layers

As the neural networks start converging, a degradation problem occurs. For plain architecures, accuracy gets saturated and then degrades rapidly. This is not due to model overfitting and implementing more complex architectures does not solve the problem. To address this issue, residual networks have been introduced in order to solve the degradation problem. Instead of fitting stacked layers, we let these layers fit a residual mapping. Formulation of this idea might be realized via implementing "shortcut connection" in a neural network. Shortcut connections are created via an identity mapping between one or more skipped layers. Details and some mathematical background can be found in [He+15].

6 Exploring data augmentations

TODO

References

[He+15] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: 7 (Dec. 2015).

[GBC] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*.

[KNH] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. "CIFAR-10 (Canadian Institute for Advanced Research)". In: (). URL: `http://www.cs.toronto.edu/~kriz/cifar.html`.