

Object Recognition and Image Understanding

Exercise Sheet 6

Shivali Dubey, Sascha Stelling

June 8, 2018

Question 1

- **Team:**

Shivali:

- Implementation
- Neural network setup

Sascha:

- Implementation
- Preparation of the dataset
- Running training phase

- **Problem Definition:** *Object detection and multiclass image classification.*

Label images from a set of labels with the assumption that each image can only get labelled by one class

- **Dataset:** Tiny ImageNet

- **Approach:** Extract features from an input image into a learned filter-bank using CNNs containing convolution layers, PReLU¹ as an activation function and pooling layers and pooling layers using max pooling to reduce the output size for the next neuron and because the exact location of a feature is less important than the rough location relative to other features. To detect features in the first convolution layer we use the Scale-Invariant Feature Transform (SIFT)² which computes the Difference of Gaussians (DoG)³. DoG is used to detect blobs by subtracting two blurred images

¹https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/He_Delving_Deep_into_ICCV_2015_paper.pdf?spm=5176.100239.blogcont55892.28.pm8zm1&file=He_Delving_Deep_into_ICCV_2015_paper.pdf

²text

³<http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/diffgaussians/index.html>

from another with different Gaussian kernels. The maxima and minima of this operation are used as key feature locations for the next neurons.

- **Evaluation & Expected Results:** Calculate training error and minimize it while maximizing accuracy. We hope to get an accuracy of at least 0.7
- **Hardware:**
- **Excluded Presentation Date:**