Lab: CudaVision – Learning Vision Systems on Graphics Cards (MA-INF 4308)

Final Assignment Topics

1.8.2016

Prof. Sven Behnke
Dr. Seongyong Koo

universitätbonn ais

In general,

Topic

- Select one of the topics descriped in the following pages
- Read and study suggested references
- Implement and train the existing models with the suggested training datasets.
- Test accuracy or equivalent performance measure with validation dataset. Try to adjust hyper-parameters of the model to achieve best performance.
- (Optional) Suggest your own model and compare its performance with existing ones.

Schedule

- 1/8-5/8: Discussion week for topic selection
- 21/9-30/9: Discussion week for lab report
- 30/9: Lab report deadline

Lab report

- Use the attached LaTeX template. The main idea is to simulate a tiny “thesis”.
- Write a nice, readable introduction, describe which methods you have used, include some math, put your results in a table, include a figure (both with caption), reference table and figure in your text when you discuss the results. Write a conclusion.
- Make sure to cite correctly. Ensure you do not copy from anywhare without citing the source. When you refer codes from open source community.
- Write an abstract that summarizes everything you wrote. Use a spell checker and/or let somebody else read your report before turnning it in.
- Your lab report should have 6-8 pages. Margins are fairly wide, so there isn't much to write. Try to be brief, but readable and informative.

Topic 1

3D convolutional logistic regression for 3D plat root reconstruction

Objectives

- Implementing 3D convolutional logistic regression for root and soil classification from MRI images
- Comparing three different representations (serialized 1D vector, KxK 2D images + K-channel, KxKxK 3D volume) in terms of training time and classification accuracy
- Developing a training method robust to imbalanced and noisy data

universitätbonn ais

Topic 1

3D convolutional logistic regression for 3D plat root reconstruction

References

- http://www.ais.uni-bonn.de/papers/CCIS_2013_Schulz.pdf
- http://www.ais.uni-bonn.de/papers/VISAPP2012_schulz.pdf
- http://www.ele.uri.edu/faculty/he/PDFfiles/ImbalancedLearning.pdf

Dataset

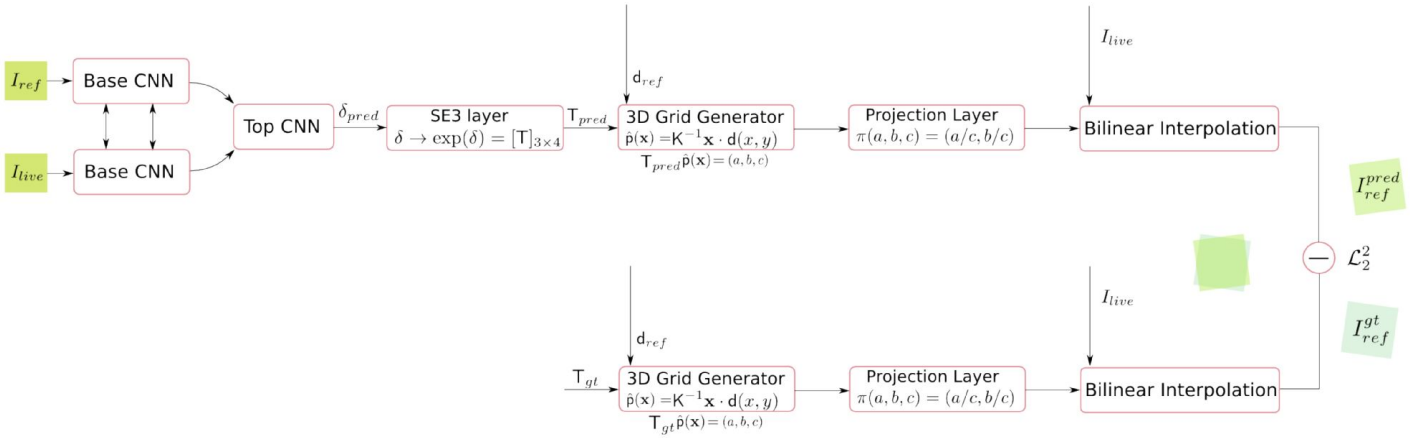
- Plant root MRI images will be provided from our lab

Topic 2

3D spatial transformer networks for visual odometry

Objectives

- Implementing SE3 Layer that was proposed in gvnn paper for TensorFlow
- It is an extention of 2D Spatial Transfer Networks in TensorFlow to SE3 space to train and estimate 6D pose of an object
- As suggested in gvnn paper, use the layer to build a network model to estimate visual odometry of an RGB-D camera.
- The model can be an alternative of existing SLAM methods. Use SLAM datasets to train the network.



universitätbonn ais

Topic 2

3D spatial transformer networks for visual odometry

References

- Spatial Transformer Networks, NIPS 2015
  http://papers.nips.cc/paper/5854-spatial-transformer-networks
  TensorFlow implementation:
  https://github.com/tensorflow/models/tree/master/transformer
- gvnn: Neural Network Library for Geometric Computer Vision
  http://arxiv.org/abs/1607.07405

Dataset

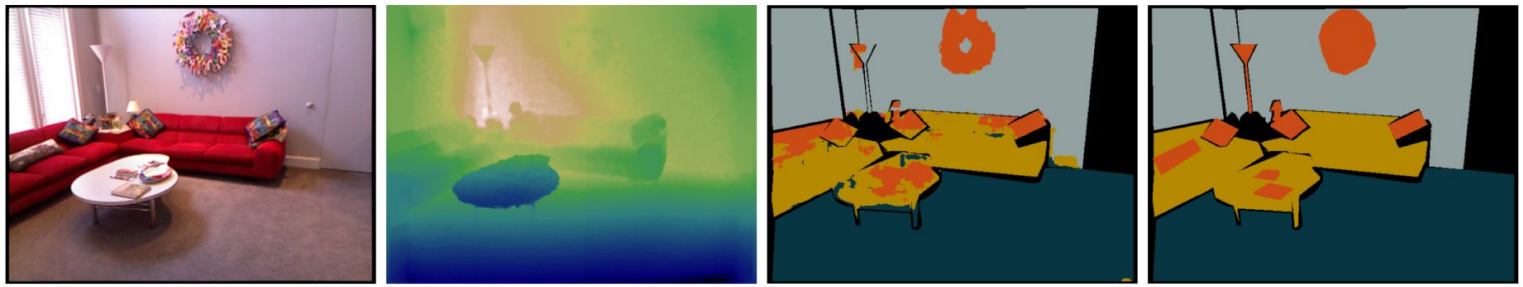
- CoRBS: Comprehensive RGB-D Benchmark for SLAM using Kinect v2
  http://corbs.dfki.uni-kl.de/
- TUM RGB-D SLAM Dataset
  http://vision.in.tum.de/data/datasets/rgbd-dataset

Topic 3

Object-class segmentation using recurrent neural networks

Objectives

- Object-class segmentation (sementic segmentation) is to find object classes in an image or a series of images in pixel level.
- Fully Convolutional Networks is one of the best models for object-class segmentation.
- When RGB-D video data is available, we could improve segmentation performance rather then using a single RGB image.
- Implement a FCN variant model with LSTM to perform object-class segmentation from RGB-D video stream.



(a) RGB frame (b) Depth (c) Prediction (d) Ground truth

Topic 3

Object-class segmentation using recurrent neural networks

References

- Fully Convolutional Networks for Semantic Segmentation, CVPR15
  http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html
  TensorFlow implementation
  https://github.com/MarvinTeichmann/tensorflow-fcn
- Recurrent Convolutional Neural Networks for Object-Class Segmentation of RGB-D Video, IJCNN 2015
  http://ais.uni-bonn.de/papers/IJCNN_2015_Pavel.pdf

Dataset

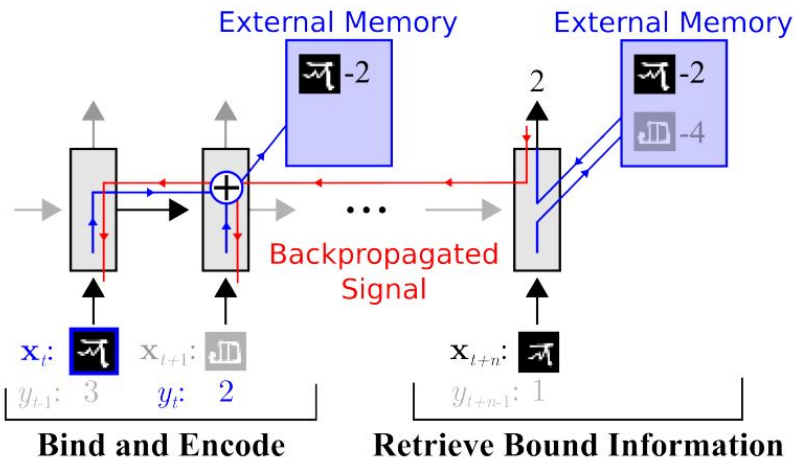
- MS COCO
  http://mscoco.org/dataset/#detections-challenge2016
- NYU dataset V2
  http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html

Topic 4

One-shot learning for image classification with memory augmented networks

Objectives

- A memory-augmented neural networks is able to rapidly assimilate new data and leverage this data to make accurate predictions after only a few samples, one-shot learning.
- The main objective is to study recent memory augmented networks and implement one-shot learning on TensorFlow.
- The one-shot accuracies of different models should be compared.



universitätbonn ais

Topic 4

One-shot learning for image classification with memory augmented networks

References

- One-shot Learning with Memory-Augmented Neural Networks
  http://arxiv.org/abs/1605.06065
  Theano implementation
  https://github.com/tristandeleu/ntm-one-shot
- Matching Networks for One Shot Learning
  http://arxiv.org/abs/1606.04080

Dataset

- Omniglot, transpose of MNIST
  https://github.com/brendenlake/omniglot
- ImageNet, ILSVRC-2012,
  http://www.image-net.org/challenges/LSVRC/2012/