Am Dienstag, 12.09.2017, 15:15 hst berichtet Herr Ahmed Osman über das Ergebnis seiner Masterarbeit:

"Dual Recurrent Attention Units for Visual Question Answering"

Abstract:
A vast field of real-world AI applications deals with inputs from different medium. After the breakthroughs in visual recognition and text classification, Visual Question Answering (VQA) have recently gained traction in the AI community. The VQA task requires the machine to answer a question in natural language regarding a given image. Multimodal reasoning tasks require knowledge from multiple disciplines, such as Computer Vision (CV) and Natural Language Processing (NLP). To solve such tasks, rich representations and appropriate architectures need to be explored.

The goal of this thesis is to develop and study Deep Neural Networks (DNNs) for VQA. Attention mechanisms, which are widely used in VQA models, guide DNNs to focus on important segments of the visual and textual data. We propose a method that utilizes recurrent layers to generate explicit visual and textual attention which we refer to as the Dual Recurrent Attention Units (DRAU) network. We believe that the memory characteristic of recurrent layers offers a rich joint embedding of visual and textual features which helps provide a more robust attention mechanism. Our model outperforms the previous first place team on the COCO-based VQA 2016 challenge and achieves near state-of-the-art performance on the VQA 1.0 dataset. Furthermore, we experiment with replacing attention mechanisms in other models with our attention implementation and show boosted performance.

Zeit: Tuesday, 12.09.2017, 15:15 hst   Achtung: Geänderte Zeit!!!
Ort: Geb. 52, 02-017

Interessenten sind herzlich eingeladen. Weitere Informationen bei:
Prof. Dr.-Ing. Thomas Brox, Tel: 0761/203-8261
Email: brox@informatik.uni-freiburg.de
http://lmb.informatik.uni-freiburg.de/lectures/oberseminar/