Zhiheng Li– Curriculum Vitae

Address 3211 Wegmans Hall, 250 Hutchison Road
Rochester, NY, USA
Email zhiheng.li@rochester.edu

(last update: Feb 2020)

Education

2018 - now **PhD**, Department of Computer Science, University of Rochester, NY, USA
Advisor Prof. Chenliang Xu

2014 - 2018 **Bachelor of Engineering**, Department of Computer Science, Wuhan University, Hubei, China
Advisor Prof. Zhenzhong Chen

Publications

Deep Grouping Model for Unified Perceptual Parsing
Zhiheng Li, Wenxuan Bao, Jiayang Zheng, Chenliang Xu
The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020

Learning a Weakly-Supervised Video Actor-Action Segmentation Model with a Wise Selection
Jie Chen, **Zhiheng Li**, Jiebo Luo, Chenliang Xu
The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020

Lip Movements Generation at a Glance
Lele Chen*, **Zhiheng Li***, Ross K Maddox, Zhiyao Duan, Chenliang Xu (* **equal contribution**)
The European Conference on Computer Vision (ECCV), 2018

Research interests

Semantic Segmentation, Action Recognition, Video Generation.

Research experience

Jul. 2019 - Present University of Rochester
Deep Grouping Model for Unified Perceptual Parsing

The perceptual-based grouping process produces a hierarchical and compositional image representation that helps both human and machine vision systems recognize heterogeneous visual concepts. Examples can be found in the classical hierarchical superpixel segmentation or image parsing works. However, the grouping process is largely overlooked in modern CNN-based image segmentation networks due to many challenges, including the inherent incompatibility between the grid-shaped CNN feature map and the irregular-shaped perceptual grouping hierarchy. Overcoming these challenges, we propose a deep grouping model (DGM) that tightly marries the two types of representations and defines a bottom-up and a top-down process for feature exchanging. When evaluating the model on the recent Broden+ dataset for the unified perceptual parsing task, it achieves state-of-the-art results while having a small computational overhead compared to other contextual-based segmentation models. Furthermore, the DGM has better interpretability compared with modern CNN methods.

Aug. 2018 - Jun. 2019 University of Rochester
Learning the Interplay between Actor and Action via Metric Learning

- Propose a metric learning-based method on joint learning of actor and action, enabling the model to learn the interplay between actor and action.
- Along with metric learning method, using an data sampling method to strengthen metric learning's performance by feeding the model pair (or triplet) of data with such property: "same actor, different actions" or "same action, different actors".
- The metric learning-based method is incorporated into both multi-task learning model and hierarchical based learning model. A2D dataset is used as the training and testing set.

Sep. 2017 - Oct. 2017 University of Rochester
Lip Movements Generation

- Worked on generating video clip of lip movements based on audio speech.
- Proposed a method that strengthens correspondence between audio information and visual information.
- The project is written in Python and is based on PyTorch.

Sep. 2017 - Oct. 2017 University of Rochester
Deformable Convolution 3D

- Refactored CUDA C++ code for extending deformable convolution to 3D model.
- Applied deformable convolution 3D in action recognition task and tested on different action recognition datasets, such as UCF101, Kinetics, etc.

Apr. 2016 - Jun. 2017 Wuhan University
Abnormal Event Detection in Surveillance Video

- Focused on the Surveillance Event Detection (SED) task of TRECVID to detect abnormal events from surveillance camera video.
- Incorporated Faster R-CNN, an advanced object detection algorithm, as a pedestrian detector into our model.
- Developed a visual tool for the purpose of showing the detection results using Qt and OpenCV.

Technical Skills

Programming Languages: Python, CUDA C/C++, Java, C, C++, Ruby, Swift, Objective-C
Frameworks: PyTorch, Caffe, MxNet, OpenCV, Qt

Related Courses

University of Rochester

- CSC 577: Advanced Topics in Computer Vision
- CSC 453: Dynamic Languages and Software Development
- CSC 400: Graduate Problem Seminar

Wuhan University

- Digital Image Processing
- Data Structure
- Calculus
- Linear Algebra
- Probability and Statistics
- Discrete Mathematics

Coursera

- Neural Network and Machine Learning