

# Zhiheng Li– Curriculum Vitae

---

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

(last update: Sep 2020)

## Education

**2018 - now** **PhD Student**, 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

### **Graph Neural Network Based Coarse-Grained Mapping Prediction**

**Zhiheng Li**, Geemi P. Wellawatte, Maghesree Chakraborty, Heta A Gandhi, Chenliang Xu, Andrew D. White  
Chemical Science, 2020

### **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 experience

**Feb. 2020 -** University of Rochester  
**Aug. 2020** *Graph Neural Network Based Coarse-Grained Mapping Prediction*

The selection of coarse-grained (CG) mapping operators is a critical step for CG molecular dynamics (MD) simulation. It is still an open question about what is optimal for this choice and there is a need for theory. The current state-of-the-art method is mapping operators manually selected by experts. In this work, we demonstrate an automated approach by viewing this problem as supervised learning where we seek to reproduce the mapping operators produced by experts. We present a graph neural network based CG mapping predictor called Deep Supervised Graph Partitioning Model (DSGPM) that treats mapping operators as a graph segmentation problem. DSGPM is trained on a novel dataset, Human-annotated Mappings (HAM), consisting of 1180 molecules with expert annotated mapping operators. HAM can be used to facilitate further research in this area. Our model uses a novel metric learning objective to produce high-quality atomic features that are used in spectral clustering. The results show that the DSGPM outperforms state-of-the-art methods in the field of graph segmentation. Finally, we find that predicted CG mapping operators indeed result in good CG MD models when used in simulation.

**Jul. 2019 -** University of Rochester

**Feb. 2020** *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 -** University of Rochester

**Jun. 2019** *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 -** University of Rochester

**Oct. 2017** *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 -** University of Rochester

**Oct. 2017** *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 -** Wuhan University

**Jun. 2017** *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