KAI CHEN

No. 1433, Cailun Road, Pudong New District, Shanghai, 201203, P.R.China
Email: kchen16@fudan.edu.cn ◇ Homepage: kaichen1998.github.io

EDUCATION

Hong Kong University of Science and Technology, HK, China — *Sep 2020 - Jun 2025 (Expected)*
(Coming) Ph.D. in **Computer Science and Engineering**

Fudan University(FDU), Shanghai, China — *Sep 2016 - Jun 2020*
B.S. in **Computer Science**, Minor in **Economics**
Overall GPA: 3.7/4.0, Major GPA: 3.91/4.0, Ranking: 3/32

University of Manchester, Manchester, UK — *Sep 2018 - Jan 2019*
Exchange student in Computer Science, advised by Dr. Tingting Mu

PUBLICATIONS

- Md Alimoor Reza, Akshay Naik, **Kai Chen**, David Crandall, "Automatic Annotation for Semantic Segmentation in Indoor Scenes," *IEEE International Conference on Intelligent Robots and Systems (IROS),2019* [pdf]

HONORS

National Scholarship for Outstanding Students (1%, by Ministry of Education of P.R.China)	Sep 2018
Joel & Ruth Spira Scholarship (1%, by Lutron Electronics)	Mar 2019
Scholarship for Outstanding Undergraduate Students (5%, by Fudan University)	Oct 2017
Outstanding undergraduate of Fudan University (10%)	May 2018 & Oct 2017
1st Prize - "ChuangQingChun" Enterprising Competition FDU Division(10%)	Feb 2018

RESEARCH EXPERIENCE

Automatic Annotation for Semantic Segmentation in Indoor Scenes — June 2019 - present
Undergraduate Research — *Advisor:Dr. Md Reza and Prof. David Crandall, IUB*

· Motivation: Expensive to get image semantic labels manually, so it's necessary to generate image semantic annotation automatically without any ground truth labels. Using together with human annotation, we hope we can get a better segmentation (e.g. FCN) model
· Idea: Structural scene understanding separates an image to 2 parts first: foreground and background. We use Mask RCNN to detect foreground objects and 3D layout segmentation estimator to capture background information separately. Then we gather this two information together using a well-defined energy function to find best annotation
· Summary our work into a paper accepted by *IROS 2019* as third author (see details above)
· Now we are trying to find a way to *'fine tuning'* our Mask RCNN detector without access to any ground truth labels to improve performance

Few Shot Image Classification with Multiple Image Views — June 2019 - present
Undergraduate Research — *Advisor: Prof. David Crandall, IUB*

· Idea: instead of seeing thousands of chair images, we think humans learn to recognize a chair by seeing a chair from multiple views and generating a 3D chair model in their memory
· Based on *Geometry-Aware Recurrent Network* [link], our model uses RGB input to generate a 3D feature tensor and updates a 3D GRU memory at each image view

- Experiments are based on CORE50 Dataset which only contains multiple view images for 50 objects, 10 classes and 5 objects for each, which can be considered as a few shot learning circumstance

Unsupervised Object Detection using Variational AutoEncoder Feb 2019 - June 2019
Undergraduate Research *Advisor:Prof. Bin Li and Prof. Xiangyang Xue, FDU*

- Idea: structural image understanding sees one object at a time and joins them together to get a big picture.
- Try to do object detection on special raw images (e.g. MINST) in a unsupervised way without ground truth.
- Based on *AIR* [link], our model uses a RNN to encode one object at a time and learns to figure out how many objects there are in this image so that we can get variable-length latent representation for different images
- Use VAE to re-generate the image to get training signal

Link Prediction on Weighted Signed Social Network July 2018 - April 2019
Undergraduate Research *Advisor:Prof. Yitong Wang, FDU*

- Focus on link prediction problems in Weighted Signed Social Network (WSN)
- Come out an algorithm called MFLG, a new network embedding algorithm based on matrix factorization. We embed each node with two vectors, one *subjective* one *objective*
- We use *random walk* to get related node pairs which aren't connected directly to combine global social network features with local features (adjacent node pairs)
- We add regularization to get correct sentiment results when two predictions have same square error (e.g. when answer is 1, choose 3 instead of -1)
- Use different algorithms' prediction as features to train a linear regression model and get a more robust model
- Summarize our work into a paper as the third author (still waiting for notification)

INTERNSHIP

SenseTime, Mobile Intelligence Group (MIG) Oct 2019 - present
Research Intern *Advisor:Dr. Wenxiu Sun, Sensetime*

- Focus on research topics related to instance segmentation and weakly supervised semantic segmentation and other pixel-level computer vision problems

Indiana University Bloomington (IUB), Computer Vision Lab June 2019 - Sep 2019
Visiting Scholar *Advisor:Prof. David Crandall, IUB*

- Global Talent Attraction Program (GTAP) Scholar of Indiana University Bloomington Computer Vision Lab
- Research in weakly supervised semantic segmentation, ego-motion video understanding, neuroscience inspired by human beings and 3D reconstruction

TECHNICAL SKILLS

Program Languages	Python, Matlab, C/C++/C#, SQL, Latex
Framework	Pytorch, Tensorflow
Language	Native in Mandarin Chinese, Fluent in English CET-4(649), CET-6(619), IELTS(6.5)