XIANGTAI LI

Github: https://github.com/lxtGH, Google Scholar
NTU, S-Lab, Singapore
Tel:87909536 ⋄ lxtpku@pku.edu.cn

EDUCATION

Peking University *September 2017 - July 2022*
Phd Degree
Major: Computer Science
Adviser: Prof.Yunhai Tong

Beijing University of Posts and Telecommunications *September 2013 - June 2017*
Joint Program of BUPT and QMUL
Bachelor Degree Overall GPA: 3.8/4.0 (3/120)
Major: Internet Of Things
Graduated Adviser: Prof.Haitao Zhang

WORK EXPERIENCE

Nanyang Technological University, Research Fellow Sep 2022 - Now
Supervised by Prof. Chen Change Loy.

- Research on General Segmentation and Detection.
- Research on Video Scene Understanding.
- Research on Data Efficient Learning (Few Shot, Open Vocabulary/Set).

INTERN EXPERIENCE

SenseTime: Auto-Driving Group(*AI Exploration Project*) Feb 2020 - May 2022
Research Intern (Mentor: Dr.Guangliang Cheng and Dr.Jianping Shi, also co-work with Dr.Jiangmiao Pang and Dr.Weiwen Zhang and Prof. Chen Change Loy.) Topic: Semantic/Instance/Panoptic Segmentation

- Research on Image/Video Semantic/Instance/Panoptic Segmentation.
- Maintain the inner segmentation tools (Spring element).
- Develop new algorithms for road scene understanding. DecoupleSegNets(ECCV-2020), BiAlignSegNet(ICIP-2021), PanopticPartFormer, Video-KNet.
- Develop new algorithms for general segmentation. PFSegNets for Aerial Segmentation (CVPR-2021), Transparent SegNets for Transparent Object Segmentation (ICCV-2021), BoundarySqueeze for instance segmentation.
- Develop new algorithms for video analysis. Temporal Pyramid Routing for Video Instance Segmentation, TPR(TPMAI-2022), TransVOD(TPMAI-2022) for Video Object Detection. Video K-Net(CVPR-2022) for Video Panoptic Segmentation.

JD Explore Academy Aug 2021 - Oct 2021
Remote Cooperation *(Mentor: Prof.Dacheng Tao and Dr.Yibo Yang) Topic: Panoptic Segmentation with Depth Prediction*

- Research on Image (Video) Panoptic Segmentation with Depth Prediction and multitask dense prediction.
- Win the ICCV-2021 Segmentation and Tracking Workshop (Video + Depth Track).

DeepMotion(XiaoMi Car), Perception Group Feb 2018 - June 2019
Research Intern (Mentor: Dr.Kuiyuan Yang) *Topic: Semantic Segmentation*

- Semantic segmentation on various standard datasets like Cityscapes, ADE20k, Pascal Context.
- Re-implement the state-of-the-art segmentation models using pytorch (FCN, PSP-net, Deeplabv3/v3+)and build the semantic segmentation code base for DeepMotion.
- Research on semantic segmentation in the automatic driving scenario. Focus on both fast speed and high accuracy segmentation result.
- Propose **new semantic segmentation methods** including GFF-net(AAAI-2020), GALD-net(BMVC-2019), Dual-GCN Net(BMVC-2019) and achieve top performance on several public benchmark datasets. (including Cityscapes, ADE20k).
- Propose **Fast and Accurate Semantic Segmentation** for internal usage SFNet(ECCV-2020).

BUPT Dec 2016 - Mar 2017
Undergraduate Intern (Mentor: Prof.Haitao Zhang/Dr.Xinchen Liu) *Topic: Large Food Image Classification Application*

- Build food classification dataset. Tried different CNN models on crawed dataset, including vgg, resnet, resnext, densenet. Build the benchmark on this dataset for fine-grained classification tasks.

PUBLICATIONS

*** means equal contribution, # means mentored students by me.**

[36] PanopticPartFormer: A Unified and Decoupled Perspective for Panoptic Part Segmentation
Xiangtai Li, Shilin Xu, Yibo Yang, Haobo Yuan, Guangliang Cheng, Yunhai Tong, Zhouchen Lin, Dacheng Tao (submit to T-PAMI-2023)

[35] Tube-Link: A Simple, Efficient, and Flexible Framework for Video Panoptic Segmentation
Xiangtai Li, Haobo Yuan, Wenwei Zhang, Guangliang Cheng, Jiangmiao Pang, Chen Change Loy (submit to CVPR 2023)

[34] Flying to Collapse: Evolving into the Neural Collapse Inspired Classifier for Class Incremental Learning
Haobo Yuan, Yibo Yang, **Xiangtai Li**, Lefei Zhang, Dacheng Tao (submit to CVPR 2023)

[33] Panoptic Video Scene Graph Generation
Jingkang Yang, Wenxuan Peng #, **Xiangtai Li**, Zujin Guo, Liangyu Chen, Bo Li, Zheng Ma, Wayne Zhang, Kaiyang Zhou, Chen Change Loy, Ziwei Liu (submit to CVPR 2023)

[32] Rethinking Mobile Block for Efficient Neural Models
Jiangning Zhang, **Xiangtai Li**, Jian Li, Liang Liu, Zhucun Xue, Boshen Zhang, Zhengkai Jiang, Tianxin Huang, Yabiao Wang, Chengjie Wang (submit to CVPR 2023)

[31] Reference Twice: A Simple and Unified Baseline For Few-Shot Instance Segmentation
Yue Han#, Jiangning Zhang, Zhucun Xue, Chao Xu, Xintian Shen, Yabiao Wang, Chengjie Wang, Yong Liu, **Xiangtai Li** (submit to CVPR 2023)

[30] Betrayed by Captions: Joint Caption Grounding and Generation for Open Vocabulary Instance Segmentation

Jianzong Wu* #, **Xiangtai Li***, Henghui Ding, Xia Li, Yunhai Tong, Guangliang Cheng, Chen Change Loy (submit to CVPR 2023)

[29] Pairing then Predicting: PairNet for Accurate Scene Graph Generation
Jinghao Wang #, Zhengyu Wen #, **Xiangtai Li**, Jingkang Yang, Zujin Guo, Ziwei Liu (submit to CVPR 2023)

[28] Neural Collapse Inspired Feature-Classifier Alignment for Few-Shot Class-Incremental Learning
Yibo Yang, Haobo Yuan, **Xiangtai Li**, Zhouchen Lin, Philip Torr, Dacheng Tao (submit to ICLR 2023)

[27] Towards Robust Referring Image Segmentation
Jianzong Wu #, **Xiangtai Li**, Xia Li, Henghui Ding, Yunhai Tong, Dacheng Tao (submit to CVPR 2023) arxiv

[26] EAFormer: Improving Vision Transformer Inspired by Evolutionary Algorithm
Jiangning Zhang, **XiangtaiLi**, Yibo Yang, Yong Liu, Dacheng Tao (submit to IJCV) arxiv

[25] Convolution-enhanced Evolving Attention Networks
Yujing Wang, Yaming Yang, Zhuo Li, Jiangang Bai, Mingliang Zhang, **Xiangtai Li**, Jing Yu, Ce Zhang, Gao Huang, Yunhai Tong (**T-PAMI-2023**)

[24] Inducing Neural Collapse in imbalanced training: Do We Really Need a Learnable Classifier at the End of Deep Neural Network?
Yibo Yang, Shixiang Chen, **Xiangtai Li**, Liang Xie, Zhouchen Lin, Dacheng Tao (**NIPS-2022**) arxiv.

[23] SFNet: Faster and Accurate Semantic Segmentation via Semantic Flow
Xiangtai Li, Jiangning Zhang, Guangliang Cheng, Kuiyuan Yang, Yibo Yang, Yunhai Tong, Dacheng Tao. (**submit to IJCV**, extension of previous work[5]) arxiv code

[22] Multi-task Learning with Multi-query Transformer For Dense Prediction
Yangyang Xu #, **Xiangtai Li**, Haobo Yuan, Yibo Yang, Jing Zhang, Yunhai Tong, Lefei Zhang, Dacheng Tao (submit to IJCV) arxiv.

[21] Fashionformer: A simple, effective and unified baseline for human fashion segmentation and recognition
Shilin Xu* #, **Xiangtai Li***, Jingbo Wang, Guangliang Cheng, Yunhai Tong, Dacheng Tao (**ECCV-2022**) arxiv. code

[20] TransVOD: End-to-end Video Object Detection with Spatial-Temporal Transformers
Qianyu Zhou* #, **Xiangtai Li***, Lu He #, Yibo Yang, Guangliang Cheng, Yunhai Tong, Liqing Zhang, Dacheng Tao. (**T-PAMI 2022**, extension of previous work[11].) arxiv. code

[19] Panoptic PartFormer: Learning a Unified Model for Panoptic Part Segmentation
Xiangtai Li, Shilin Xu, Yibo Yang, Guangliang Cheng, Yunhai Tong, Dacheng Tao (**ECCV-2022**) arxiv. code

[18] PolyphonicFormer: Unified Query Learning for Depth-aware Video Panopic Segmentation
Haobo Yuan* #, **Xiangtai Li***, Yibo Yang, Yunhai Tong, Dacheng Tao (**ECCV-2022**) arxiv. code

[17] Improving Video Instance Segmentation via Temporal Pyramid Routing.
Xiangtai Li, Hao He #, Yibo Yang, Kuiyuan Yang, Guangliang Cheng, Dacheng Tao (**T-PAMI-2022**) arxiv code.

[16] BoundarySqueeze: Image Segmentation as Boundary Squeezing
Hao He* #, **Xiangtai Li***, Yibo Yang, Guangliang Cheng, Yunhai Tong, Shiming Xiang, DaCheng Tao (**submit to IJCV**) arxiv. code.

[15] Video K-Net: A Simple, Strong, and Unified Baseline For Video Segmentation
Xiangtai Li, Wenwei Zhang, Jiangmiao Pang, Kai Chen, Guangliang Cheng, Yunhai Tong, Chen Change Loy (**CVPR-2022**, oral (4%)) arxiv. code

[14] Enhanced Boundary Learning for Glass-like Object Segmentation.
Hao He* #, **Xiangtai Li***, Guangliang Cheng, Jianping Shi, Yunhai Tong, Shiming Xiang **(ICCV-2021) arxiv code**

[13] Towards Efficient Scene Understanding via Squeeze Reasoning.
Xiangtai Li, Xia Li, Ansheng You, Li Zhang, Guangliang Cheng, Kuiyuan Yang, Yunhai Tong, Zhouchen Lin (**IEEE-TIP-2021**). arxiv. code.

[12] Global Aggregation then Local Distribution for Scene Parsing.
Xiangtai Li, Li Zhang, Guangliang Cheng, Kuiyuan Yang, Yunhai Tong, Xiatian Zhu, Tao Xiang (**IEEE-TIP-2021**) arxiv code

[11] End-to-End Video Object Detection with Spatial-Temporal Transformers
Lu He* #, Qianyu Zhou* #, **Xiangtai Li***, Li Niu, Guangliang Cheng, Xiao Li, Wenxuan Liu, Yunhai Tong, Lizhuang Ma, Liqing Zhang (**ACM-MM-2021**) arxiv code.

[10] Dynamic Dual Sampling Module for Fine-Grained Semantic Segmentation
Chen Shi* #, **Xiangtai Li***, Yanran Wu, Yunhai Tong, Yi Xu (**ICIP-2021**) arxiv. code.

[9] Fast and Accurate Scene Parsing via Bi-direction Alignment Networks
Yanran Wu* #, **Xiangtai Li***, Chen Shi, Yunhai Tong, Yang Hua, Tao Song, Ruhui Ma, Haibing Guan (**ICIP-2021**) arxiv. code.

[8] PointFlow: Flowing Semantics Through Points for Aerial Image Segmentation.
Xiangtai Li*, Hao He*, Xia Li, Duo Li, Guangliang Cheng, Jianping Shi, Lubin Weng, Yunhai Tong, and Zhouchen Lin (**CVPR-2021**) arxiv. code.

[7] Inverting the Inherence of Convolution for Visual Recognition
Duo Li, Jie Hu, Changhu Wang, **Xiangtai Li**, Qi She, Lei Zhu, Tong Zhang, Qifeng Chen (**CVPR-2021**) arxiv. code

[6] Improving Semantic Segmentation via Decoupled Body and Edge Supervision
Xiangtai Li, Xia Li, Li Zhang, Guangliang Cheng, Jianping Shi, Zhouchen Lin, Shaohua Tan, Yunhai Tong (**ECCV-2020**). arxiv. code

[5] Semantic Flow for Fast and Accurate Scene Parsing.
Xiangtai Li, Ansheng You, Zhen Zhu, Houlong Zhao, Maoke Yang, Kuiyuan Yang, Yunhai Tong (**ECCV-2020, oral(2%)**). arxiv. code.

[4] GFF: Gated Fully Fusion for Semantic Segmentation
Xiangtai Li, Houlong Zhao, Lei Han, Yunhai Tong, Shaohua Tan, Kuiyuan Yang (**AAAI-2020, oral(4%)**) arxiv. code.

[3] Global Aggregation then Local Distribution for Scene Understanding
Xiangtai Li, Li Zhang, Ansheng You, Maoke Yang, Kuiyuan Yang, Yunhai Tong (**BMVC-2019**) arxiv. code.

[2] Dual Graph Convolutional Network for Semantic Segmentation
Li Zhang*, **Xiangtai Li***, Anurag Arnab, Kuiyuan Yang, Yunhai Tong, Philip H. S. Torr (**BMVC-2019**) arxiv. code.

[1] Flo2seg: Motion-Aided Semantic Segmentation
Xiangtai Li, Jiangang Bai, Yunhai Tong, Kuiyuan Yang (**ICANN-2019** Long Paper).

TECHNICAL STRENGTHS

Computer Languages	python, C/C++, java
Software & Tools	pytorch, detectron2/mmdet/mmseg/mmtrack, Tensorflow, Latex, CUDA

HORNOR

Winner of Segmenting and Tracking Every Point and Pixel: 6th Workshop on ICCV-2021 Track2 (Project Leader and First Author, 2021-10-17)

National Scholarship, Ministry of Education of China (2020-2021) (Top 1% in PKU)

Haiying Star Scholarship of PKU(2020-2021) (Top 1% in PKU)

President Scholarship of PKU (2020-2021) (Top 1% in PKU)

National Scholarship, Ministry of Education of China (2019-2020) (Top 1% in PKU)

Merit Student of PKU (2019, 2020, 2021) (Top %2)

Scientific Research Award of PKU (2019, 2020, 2021) (Top %2 in PKU)

Second Prize of Excellent Undergraduate Scholarship of BUPT (2014, 2015, 2016) and PKU (2018).

Merit Student of BUPT (2014, 2015, 2016) with the Second Prize Scholarship.

2015 and 2016 MCM/ICM Contest: Meritorious Winner (top 5%)

2017 and 2022 Beijing Excellent Graduates (top 1%)

ACADEMIC PARTS

Invited talk on Panoptic Segmentation and Beyond in Baidu PaddlePaddle Group (2022.5)

Invited talk on DiDi Auto-Driving Group (2021.11)

Invited talk on HuaWei Noah Auto-Driving Group (2021.10)

Invited talk on Aligned Segmentation in JD AI research (2021.8)

Presenter of Segmenting and Tracking Every Point and Pixel Workshop at ICCV-2021 Track2: Video Segmentation + Depth (2021.10)

Invited talk on Shanghai lab (Aligned Approach for Segmentation) (2021.8)

Invited talk on ECCV-2020-oral-session in Jiangmen (2020.8)

Regular Reviewer for CVPR, ICCV, ECCV, ICLR, NIPS, ICML, AAAI, ACM-MM, T-PAMI, IEEE-TIP, IJCV, IEEE-IROS, ICIP

OPEN-SOURCE PROJECTS: LINK

I will try my best to open-source our works and related tools.

GALD-net: source code of high performance Cityscape models.

Octave-conv: recent various convolution operation on ImageNet.

DecoupleSegNet tools

FastSeg tools

ABOUT RESEARCH

My research lies in Computer Vision, with a main focus on **Segmentation** tasks during my PhD study. I am familiar with image semantic segmentation and I work in this direction for two years. I did several works on semantic segmentation and related directions including transparent objects, aerial segmentation, video domain problems. Then my main research interest turns on **instance level and full scene** segmentation tasks including instance segmentation, Panoptic segmentation and **video instance level understanding task** including Video Object Detection, Video Instance Segmentation and Video Panoptic Segmentation.

After PhD, I would try more general and diverse directions. In particular, I am interested in these directions:

- Unified Image/Video Scene Understanding, including segmentation and tracking, domain generation.
- Deep Learning with fewer examples and labels and its application.
- Video Temporal Modeling and Beyond, including forecasting, relation detection.
- Learning with multi-modal inputs, including open-vocabulary modeling.
- Joint 3D and 4D scene understanding.

MENTORS AND COLLABORATORS

I am very lucky and happy to cooperate with several amazing mentors and friends, including Dr.Kuiyuan Yang (XiaoMi Car), Prof.Li Zhang (Fudan University), Prof.Zhouchen Lin (PKU), Prof.Dacheng Tao (JD), Dr.Guangliang Cheng (SenseTime), Dr.Shijian Ping(SenseTime), Dr.Weidi Xie (Oxford). I am also very lucky to cooperate with several friends, including Dr.Xia Li (ETH), Mr.Ansheng You (Ali-DAMO), Dr.Zhen Zhu (HUST), Dr.Jiangmiao Pang (MMlab@CHUK) and Dr.Henghui Ding(NTU).

In particular, many thanks to Dr.Guangliang Cheng and Dr.Jianping Shi for enough freedom and resources to carry out research.

I have deeply mentored and co-worked with several PhD and master students, including Mr.Hao He(NLPR-CASIA, Now PhD in MMLab@CHUK), Mr.Haobo Yuan(WHU), Mr.Shilin Xu(PKU), Mr.Qianyu Zhou(SJTU), Mr.Lu He(SJTU), Ms.Yanran Wu(SJTU), Mr.Shi Chen(SJTU), Mr.Yangyang Xu(WHU), Mr.JianZong Wu(PKU) for vision research during my last two years of PhD. Thanks for their works!