

Zilong Huang

✉ speedinghzl@gmail.com

🌐 [speedinghzl](#)

🏠 [Home Page](#)

EDUCATION

Huazhong University of Science and Technology
Ph.D student in Computer Vision and Machine Learning
Advisor: Wenyu Liu, Xinggang Wang

Wuhan, China
2015.9–2020.12

University of Illinois at Urbana-Champaign
Joint-training Ph.D student in Computer Vision and Machine Learning
Advisor: Thomas Huang, Yunchao Wei, Humphrey Shi

Urbana, USA
2018.10–2019.10

Huazhong University of Science and Technology
B.S. in Electronic Information Engineering
Dian Group, Seed Class

Wuhan, China
2011.9–2015.6

RESEARCH INTERESTS

Multi-modal learning, Image/video understanding and generation, Efficient network design.

EXPERIENCES

Bytedance
Research Scientist

Singapore
Apr. 2023 - Now

- conduct cutting-edge research of Large Vision Model and Multimodal Model.

Tencent
Senior Researcher

Shanghai, China
Jan. 2021 - Mar. 2023

- conduct cutting-edge research: Mobile Transformer(CVPR22, ICLR23), NeRF(NeurIPS22), Text-to-Motion(CVPR 2023), DPT Pretraining (WACV 2024 oral)...
- ship algorithms into productions: design mobile transformers as backbone for Virtual Human Technology, which has been deployed into Super QQ show, QQ Virtual Live and QQ video call.
- Two champions of Monocular Depth Estimation and Short-video Face Parsing Challenges.

PUBLICATIONS

Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data. CVPR 2024
Lihe Yang, Bingyi Kang, Zilong Huang, Xiaogang Xu, Jiashi Feng, Hengshuang Zhao.
Most popular depth estimation model, More than 4K stars on Github.

Harnessing Diffusion Models for Visual Perception with Meta Prompts. ArXiv 2024.
Qiang Wan, Zilong Huang, Bingyi Kang, Jiashi Feng, Li Zhang.

Disentangled Pre-training for Image Matting. WACV 2024.
Yanda Li, Zilong Huang, Gang Yu, Ling Chen, Yunchao Wei, Jianbo Jiao.

Executing your Commands via Motion Diffusion in Latent Space. CVPR 2023.
Xin Chen, Biao Jiang*, Wen Liu, Zilong Huang, Bin Fu, Tao Chen, Jingyi Yu, Gang Yu.*

SeaFormer: Squeeze-enhanced Axial Transformer for Mobile Semantic Segmentation. ICLR 2023.

Qiang Wan, Zilong Huang, Jiachen Lu, Gang Yu, Li Zhang.

DCNet: Large-scale Point Cloud Semantic Segmentation with Discriminative and Efficient Feature Aggregation. TCSVT 2023.

Fukun Yin, Zilong Huang, Tao Chen, Guozhong Luo, Gang Yu, Bin Fu.

Coordinates are not lonely-Codebook Prior Helps Implicit Neural 3D Representations. NeurIPS 2022.

Fukun Yin, Wen Liu, Zilong Huang, Pei Cheng, Tao Chen, Gang Yu.

TopFormer: Token Pyramid Transformer for Mobile Semantic Segmentation. CVPR 2022.

Wenqiang Zhang, Zilong Huang*, Gang Yu, Tao Chen, Guozhong Luo, Xinggang Wang, Wenyu Liu, Chunhua Shen.*

Shuffle Transformer: Rethinking Spatial Shuffle for Vision Transformer. ArXiv 2021.

Zilong Huang, Youcheng Ben, Guozhong Luo, Pei Cheng, Gang Yu, Bin Fu.

Shuffle Transformer with Feature Alignment for Video Face Parsing. CVPRW 2021.

Rui Zhang, Yang Han, Zilong Huang, Pei Cheng, Guozhong Luo, Gang Yu, Bin Fu.

Ranked No. 1 of Short-video Face Parsing Track of The 3rd Person in Context (PIC) Challenge 2021.

A Simple Baseline for Fast and Accurate Depth Estimation on Mobile Devices. CVPRW 2021.

Ziyu Zhang, Yicheng Wang, Zilong Huang, Guozhong Luo, Gang Yu, Bin Fu.

Ranked No. 1 of MAI Monocular Depth Estimation Challenge 2021.

Human De-occlusion: Invisible Perception and Recovery for Humans. CVPR 2021.

Qiang Zhou, Shiyin Wang, Yitong Wang, Zilong Huang, Xinggang Wang.

Half-Real Half-Fake Distillation for Class-Incremental Semantic Segmentation. ArXiv 2021.

Zilong Huang, Wentian Hao, Xinggang Wang, Mingyuan Tao, Jianqiang Huang, Wenyu Liu, XianSheng Hua.

High-Resolution Deep Image Matting. AAAI 2021.

Haichao Yu, Ning Xu, Zilong Huang, Yuqian Zhou, Humphrey Shi.

AlignSeg: Feature-Aligned Segmentation Networks. TPAMI.

Zilong Huang, Yunchao Wei, Xinggang Wang, Wenyu Liu, Thomas Huang, Humphrey Shi.

Deep learning-based automated image segmentation for concrete petrographic analysis. CCR 2020.

Yu Song, Zilong Huang, Chuanyue Shen, Humphrey Shi, David A Lange.

CCNet: Criss-Cross Attention for Semantic Segmentation. TPAMI.

Zilong Huang, Xinggang Wang, Yunchao Wei, Lichao Huang, Humphrey Shi, Wenyu Liu, Thomas Huang.

Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. CVPR 2020

Mang Tik Chiu, Xingqian Xu, Yunchao Wei, Zilong Huang, Alexander Schwing, Robert Brunner, Hrnat Khachatryan, Hovnatan Karapetyan, Ivan Dozier, Greg Rose, David Wilson, Adrian Tudor, Naira Hovakimyan, Thomas S Huang, Humphrey Shi

Motion-Guided Spatial Time Attention for Video Object Segmentation. ICCVW 2019.

Qiang Zhou, Zilong Huang, Lichao Huang, Yongchao Gong, Han Shen, Wenyu Liu, Xinggang Wang.

Ranked No. 2 of Youtube Video Object Segmentation Challenge 2019.

Semantic Image Segmentation by Scale-Adaptive Networks. TIP 2019.

Zilong Huang, Chunyu Wang, Xinggang Wang, Wenyu Liu, Jingdong Wang.

CCNet: Criss-Cross Attention for Semantic Segmentation. ICCV 2019.

Zilong Huang, Xinggang Wang, Lichao Huang, Chang Huang, Yunchao Wei, Wenyu Liu.

More than 2700 citations, 1.4k stars on Github. PaperDigest Most Influential ICCV 2019 papers (5th-Place). Applications of CCNet also include AlphaFold2.

SPGNet: Semantic Prediction Guidance for Scene Parsing. ICCV 2019.

Bowen Cheng, Liang-Chieh Chen, Yunchao Wei, Yukun Zhu, Zilong Huang, Jinjun Xiong, Thomas Huang, Wen-Mei Hwu, Humphrey Shi.

Devil in the Details: Towards Accurate Single and Multiple Human Parsing. AAAI 2019.

Tao Ruan, Ting Liu*, Zilong Huang, Yunchao Wei, Shikui Wei, Yao Zhao, Thomas Huang.*

Ranked No. 1 on all human parsing tracks in the 2nd LIP Challenge 2018.

Proposal, Tracking & Segmentation: A Cascaded Network for Video Object Segmentation. ECCVW 2018.

Zilong Huang, Qiang Zhou*, Xinggang Wang, Yongchao Gong, Han Shen, Lichao Huang, Chang Huang, Wenyu Liu.*

Ranked No. 2 of Youtube Video Object Segmentation Challenge 2018.

A PyTorch Semantic Segmentation Toolbox. Technical report 2018.

Zilong Huang, Yunchao Wei, Xinggang Wang, Wenyu Liu.

Weakly-supervised semantic segmentation network with deep seeded region growing. CVPR 2018.

Zilong Huang, Xinggang Wang, Jiasi Wang, Wenyu Liu, Jingdong Wang.

Object-level proposals. ICCV 2017.

Jianxiang Ma, Anlong Ming, Zilong Huang, Xinggang Wang, Yu Zhou.

Deep patch learning for weakly supervised object classification and discovery. PR 2017.

Peng Tang, Xinggang Wang, Zilong Huang, Xiang Bai, Wenyu Liu.

ACADEMIC ACTIVITIES

- Talk at Fudan University: "TopFormer: Token Pyramid Transformer for Mobile Semantic Segmentation", 2022.
- CVPR21 the 3rd Person in Context (PIC) Workshop talk: "Shuffle Transformer with Feature Alignment for Video Face Parsing", 2021.
- ECCV18 Large-scale Video Object Segmentation Challenge Workshop talk: "Proposal Tracking and Segmentation (PTS): A cascaded network for video object segmentation", 2018.
- VALSE Pixel Level Image Understanding Workshop talk: "Weakly-Supervised Semantic Segmentation Network with Deep Seeded Region Growing", 2018.
- Reviewer for the following journals/conferences: TPAMI/IJCV/TIP/TNNLS/PR/TCSVT/Neurocomputing/CVPR/ICCV/ ECCV/ICLR/ACCV/WACV/AAAI.

AWARDS AND ACHIEVEMENTS

- World's Top2 Scientists, Stanford University/Elsevier, 2023..
- Outstanding Contributor for shipping Mobile Transformer into the production, Tencent, 2022.
- Excellent Doctoral Dissertation Award, China Society of Image and Graphics (CSIG) 2021.
- Global Top 100 Chinese Rising Stars in Artificial Intelligence, Baidu, 2021.
- 1st Place in the MAI Monocular Depth Estimation Challenge in conjunction with CVPR 2021.
- 1st Place in the Short-video Face Parsing Track of The 3rd Person in Context (PIC) Workshop in conjunction with CVPR 2021.
- 2nd Place in the 2st Large-scale Video Object Segmentation Challenge Workshop in conjunction with ICCV 2019.
- 2nd Place in the 1st Large-scale Video Object Segmentation Challenge Workshop in conjunction with ECCV 2018.
- National Ph.D. Scholarship, Ministry of Education, China, 2019.
- Zhixing Graduate Student Scholarship, HUST, 2019.
- Outstanding Graduate, HUST, 2015.

SKILLS

- **Programming:** Reasonably familiar with Python, Java, C/C++ and CUDA. Good spirit of open source with **>8K stars in Github**.
- **Computer vision and machine learning:** Familiar with several CV/ML algorithms including semantic segmentation, object detection/recognition, pose estimation, video object segmentation and convolutional neural networks. Proficient in deep learning framework Caffe and Pytorch, as well as general purpose libraries including OpenCV, scikit-learn and VLFeat.
- **Natural languages:** Mandarin (Native); English (Familiar).