Yushan Zhang

+46-700418409 (+86-15624950396) | yushanzhang324@gmail.com

profile | ZhangYushan3

Linköping, Sweden

Education

- **Linköping University** *10 2021 - now*
  Ph.D in Computer Vision Linköping, Sweden
  - Supervisor: Michael Felsberg
  - Research interest: 3D Computer Vision in general, including 3D point cloud perception, Scene flow estimation, Computer vision for robotics, 3D scene understanding, Open-vocabulary object segmentation and tracking, Diffusion models, Vision-language models.

- **Beijing Institute of Technology** *09 2018 - 06 2021*
  M.S. in Optical Engineering Beijing, China
  - Grade: 3.8/4.0

- **Beijing Institute of Technology** *09 2014 - 06 2018*
  B.S. in Electronic Engineering Beijing, China
  - GPA: 3.9/4.0, Ranking: 1/30

Experience

- **NVIDIA** *06 2024 - 11 2024*
  Research Intern in Computer Vision Remote
  - Supervisor: Laura Leal-Taixe
  - Work on the SAL project (Towards Learning to Segment Anything in LiDAR), extending it to 4D, i.e. segment and track anything in LiDAR. By leveraging state-of-the-art foundation models SAMv2, CLIP, we manage to craft the first open-vocabulary 4D LiDAR panoptic segmentation for autonomous driving scenes.

- **Linköping University** *10 2021 - now*
  Teaching Assistant Linköping, Sweden
  - Teached Courses: Deep Learning; Computer Vision for Video Analysis; Signals and Images; Images and Graphics Project Course; Medical Images; Signal and Image Processing

Publications

[1] **Zhang Y**, Osep A, Leal-Taixé L, et al. (2025). **Zero-Shot 4D Lidar Panoptic Segmentation**. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (**CVPR 2025**)*. CCF-A

We propose SAL-4D (Segment Anything in Lidar-4D), a method that utilizes multi-modal sensory robotic setups as a bridge to distill recent developments in Video Object Segmentation (VOS) in conjunction with off-the-shelf Vision-Language foundation models to Lidar. We utilize VOS models to pseudo-label tracklets in short video sequences, annotate these tracklets with sequence-level CLIP tokens, and lift them to the 4D Lidar space using calibrated multi-modal sensory setups to distill them to our SAL-4D model.

[2] **Zhang Y**, Wandt B, Magnusson M, et al. (2024). **DiffSF: Diffusion Models for Scene Flow Estimation**. *Advances in Neural Information Processing Systems, 37. (**NeurIPS 2024** **Spotlight**)*. CCF-A

1. We introduce DiffSF, leveraging diffusion models to solve the full scene flow estimation problem, where the inherent noisy property of the diffusion process filters out noisy data, thus, increasing the focus on learning the relevant patterns. 2. DiffSF introduces randomness to the scene flow estimation task, which allows us to predict the uncertainty of the estimates without being explicitly trained for this purpose. 3. We develop a novel architecture that combines transformers and diffusion models for the task of scene flow estimation, improving both accuracy and robustness for a variety of datasets.

Project Page: [] Code: []

[3] **Zhang Y**, Edstedt J, Wandt B, et al. (2023). **GMSF: Global Matching Scene Flow**. *Advances in Neural Information Processing Systems, 36. (**NeurIPS 2023**)*. CCF-A

1. A hybrid local-global-cross transformer architecture is introduced to learn accurate and robust feature representations of 3D point clouds. 2. Based on the similarity of the hybrid features, we propose to use a global matching process to solve the scene flow estimation. 3. Extensive experiments on popular datasets show that the proposed method outperforms previous scene flow methods by a large margin.

Project Page: [🌐] Code: [GitHub]

[4] Jonnarth A, **Zhang Y**, Felsberg M. (2024). **High-fidelity Pseudo-labels for Boosting Weakly-Supervised Segmentation**. *Winter Conference on Applications of Computer Vision. 2024: 1010-1019. (WACV 2024)*. CCF-B

[5] **Zhang Y**, Robinson A, Magnusson M, et al. (2023). **Leveraging Optical Flow Features for Higher Generalization Power in Video Object Segmentation**. *IEEE International Conference on Image Processing. (ICIP 2023)*. CCF-C

[6] Kristan M, et al. (2023). **The first visual object tracking segmentation vots2023 challenge results**. *Proceedings of the IEEE/CVF International Conference on Computer Vision. (ICCV 2023)*. CCF-A

[7] Kristan M, et al. (2022). **The tenth visual object tracking vot2022 challenge results**. *European Conference on Computer Vision. (ECCV 2022)*. CCF-A

[8] L Wu, T Xu, **Zhang Y**, et al. (2021). **Multi-channel feature dimension adaption for correlation tracking**. *IEEE Access 9, 63814-63824*.

[9] **Zhang Y**, J Li, F Wu, et al. (2020). **Background perception for correlation filter tracker**. *EURASIP Journal on Wireless Communications and Networking 2020, 1-12*.

[10] F Wu, T Xu, L Wu, **Zhang Y**, et al. (2020). **Adaptive and complementary correlation filter with dynamic contextual constraints**. *IEEE Access 8, 141895-141909*.

[11] Kristan M, et al. (2020). **The eighth visual object tracking VOT2020 challenge results**. *European Conference on Computer Vision. (ECCV 2020)*. CCF-A

[12] **Zhang Y**, T Xu. (2017). **A research of road centerline extraction algorithm from high resolution remote sensing images**. *Applications of Digital Image Processing XL 10396, 375-385*.

Skills

- **Programming Languages:** Python, Matlab, C++, C, C#
- **Languages:** Chinese (native), English (fluent), Swedish (beginner)

Honors and Awards

- **Beijing Excellent Graduate Title (Top 2%)** *06 2018*
- **Excellent Graduate Title of BIT (Top 2%)** *06 2018*
- **Outstanding Student Title for Undergraduate Students of BIT (3 times) (Top 5%)** *2014 - 2018*
- **Outstanding Student Scholarship for Undergraduate Students of BIT (7 times) (Top 5%)** *2014 - 2018*
- **Second Prize of College Students' Physics Competition in Beijing (Top 2%)** *12 2016*
- **Chinese National Scholarship (Top 2%)** *2015 - 2016*
- **Chinese National Scholarship (Top 2%)** *2014 - 2015*

Other Experience

- **Conference Reviewer**
  CVPR 2025, ICCV 2025
- **VOT Challenge Technical Committee**
  VOT2023, VOT2022, VOT2020 [🌐]
- **SSBA/SSDL 2024** *2024*
  Oral Session Chair [🌐]