

Zhimin Chen

+1(858) 319-6270 zhiminc@clemson.edu

◆ [Homepage](#) ◆ [Google Scholar](#) ◆ [GitHub](#)

EDUCATION

Clemson University

Ph.D. in Automotive Engineering
Advisor: Bing Li

Aug. 2020 - Dec. 2024

The Ohio State University

Master of Electrical and Computer Engineering

Aug. 2018 - May. 2020

Northeastern University, China

Bachelor of Electronic Engineering

Sep. 2014 - Jul. 2018

RESEARCH INTERESTS

I am primarily focused on studying computer vision and deep learning, with a particular emphasis on self-supervised learning, semi-supervised learning, 3D detection/segmentation, 3D zero-shot learning, multi-modality learning, foundation models, and video generation.

Technical Skills

- **Programming:** Python, C++, Linux/Unix, CUDA
- **Languages:** English, Mandarin

RESEARCH EXPERIENCE

Amazon, Applied Scientist Intern

May. 2024 – Oct. 2024

Image quality assessment via foundation models.

- Design an advanced self-supervised learning pre-training strategy to enhance the extraction of quality features with foundation models. Submitted to CVPR 2025.

Honda Research Institute, Research Intern

Jan. 2024 – May. 2024

3D occupancy forecasting via generative cellular automata.

- Designed and developed an advanced 3D occupancy forecasting method using the concept of generative cellular automata.
- Specialized in the progressive generation, focusing on a sparse set of occupied voxels and their neighborhoods, enabling an expressive sparse convolutional network.

Clemson University, Research Assistant

Aug. 2020 – Now

Text to Game Effects Generation via Video Diffusion

- Create a video game effects dataset with corresponding text description. Propose a novel structure for text to game effects video generation. Submitted to CVPR 2025.

Zero-shot 3D semantic segmentation via learning from zero-shot labels.

- Proposed a novel approach to train the network with noise annotations obtained from existing zero-shot methods. Submitted to WACV 2025.

3D Self-supervised learning via 3D to multi-view masked autoencoder.

- Proposed a 3D-to-2D generative pre-training method based on masked autoencoders to fully leverage the inherent multi-view properties of 3D data. Submitted to ICLR 2025.

Distillation of knowledge from 2D foundation models to 3D domains.

- Proposed novel tokenization method, ground-balanced contrastive learning, and two-stage frame for 3D multi-modality learning via foundation models. Published on **NeurIPS 2024**.

Multi-modal self-supervised learning based on foundation models.

- Proposed a self-supervised 3D method named Bridge3D for 3D scene understanding based on multiple foundation models and masked autoencoders. Published on **NeurIPS 2023**.

Class-level confidence based 3D semi-supervised learning.

- Proposed a novel dynamic thresholding and resampling strategy based on class-level confidence for 3D semi-supervised learning. Published on **WACV 2023**.

Prototype-based multi-modal semi-supervised learning.

- Proposed a novel multimodal semi-supervised learning method by introducing consistency constraint and a multimodal contrastive prototype loss. Published on **BMVC 2021**.

Publications

Preprints

- **Zhimin Chen**, Bing Li " Zero-shot 3D Semantic Segmentation via Learning from Zero-shot Labels". (2024). Submitted to WACV 2025.
- **Zhimin Chen**, Yingwei Li, Longlong Jing, Liang Yang, Bing Li " Point Cloud Self-supervised Learning via 3D to Multi-view Masked Learner." (2024). Submitted to ICLR 2025.
- Longlong Jing, **Zhimin Chen**, Bing Li, and Yingli Tian. "Self-Supervised Modality-Invariant and Modality-Specific Feature Learning for 3D Objects." (2021).

Publications

- **Zhimin Chen**, Bing Li. " SAM-Guided Masked Token Prediction for 3D Scene Understanding. "Advances in neural information processing systems (NeurIPS) 2024.
- **Zhimin Chen**, Longlong Jing, Yingwei Li, and Bing Li. "Bridging the Domain Gap: Self-Supervised 3D Scene Understanding with Foundation Models." Advances in neural information processing systems (NeurIPS). 2023.
- **Zhimin Chen**, Longlong Jing, Liang Yang, Yingwei Li, and Bing Li. "Class-Level Confidence Based 3D Semi-Supervised Learning." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 633-642. 2023.
- **Zhimin Chen**, Longlong Jing, Yang Liang, Yingli Tian, Bing Li "Multimodal Semi-Supervised Learning for 3D Objects." The British Machine Vision Conference (BMVC). 2021.