

Zhimin Chen

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

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

EDUCATION

Clemson University Ph.D. in Automotive Engineering Advisor: Bing Li	Aug. 2020 - May. 2025
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 3D self-supervised learning [1, 4, 5], 3D semi-supervised learning [6, 7], 3D detection/segmentation [1, 2, 3, 5, 6], 3D zero-shot learning [2], and multi-modality learning [1, 3, 4, 5, 7]. I am also interested in exploring the utilization of foundational models in 3D areas [1, 5].

Technical Skills

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

RESEARCH EXPERIENCE

Amazon, Applied Scientist Intern May. 2024 – Current

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.
- Achieve SOTA performance on popular image quality assessment datasets.

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.
- Achieved competitive results compared to concurrent 3D occupancy prediction works.

Clemson University, Research Assistant Aug. 2020 – Jan. 2024

Distillation of knowledge from 2D foundation models to 3D domains [1].

- Proposed a 2D-to-3D knowledge distillation framework with the guidance of SAM masks.

Zero-shot 3D semantic segmentation via learning from zero-shot labels [2].

- Proposed a novel approach to train the network with noise annotations obtained from existing zero-shot methods.
- Achieved SOTA performance on ScanNet in 3D zero-shot semantic segmentation task.

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

- Proposed a 3D-to-2D generative pre-training method based on masked autoencoders to fully leverage the inherent multi-view properties of 3D data.
- Achieved SOTA performance in ModelNet40 and ScanObjectNN dataset in 3D object detection and classification tasks.

Multi-modal self-supervised learning based on foundation models [5].

- Proposed a self-supervised 3D method named Bridge3D for 3D scene understanding based on multiple foundation models and masked autoencoders.
- Outperformed SOTA methods by at most 10.3% in ScanNet, SUN RGB-D, and S3DIS datasets in both 3D object detection and segmentation tasks.

Class-level confidence based 3D semi-supervised learning [6].

- Proposed a novel dynamic thresholding and resampling strategy based on class-level confidence for 3D semi-supervised learning.
- Outperformed the SOTA method by at most 5.9% in ModelNet40, ScanObjectNN, ScanNet, and SUN RGB-D datasets.

Prototype-based multi-modal semi-supervised learning [7].

- Proposed a novel multimodal semi-supervised learning method by introducing consistency constraint and a multimodal contrastive prototype loss.
- Outperformed SOTA methods in the ModelNet40 dataset by at most 4.7%.

Publications

Preprints

- [1] **Zhimin Chen**, Bing Li. " SAM-Guided Masked Token Prediction for 3D Scene Understanding. " (2024). In Submission.
- [2] **Zhimin Chen**, Yingwei Li, Longlong Jing, Liang Yang, Bing Li " Zero-shot 3D Semantic Segmentation via Learning from Zero-shot Labels". (2024). In Submission.
- [3] **Zhimin Chen**, Yingwei Li, Longlong Jing, Liang Yang, Bing Li " Point Cloud Self-supervised Learning via 3D to Multi-view Masked Autoencoder." (2024). In Submission.
- [4] Longlong Jing, **Zhimin Chen**, Bing Li, and Yingli Tian. "Self-Supervised Modality-Invariant and Modality-Specific Feature Learning for 3D Objects." (2021).

Publications

- [5] **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.
- [6] **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.
- [7] **Zhimin Chen**, Longlong Jing, Yang Liang, Yingli Tian, Bing Li "Multimodal Semi-Supervised Learning for 3D Objects." The British Machine Vision Conference (BMVC). 2021.