Jialei Chen
Research Internship Application

jialei_chen2021@163.com
psmobile.github.io

名古屋大学
NAGOYA UNIVERSITY

EDUCATION

Ph.D student, Intelligent Systems (Ph.D advisor: Daisuke Deguchi, Hiroshi Murase)
School of Informatics, Nagoya University   **2022.10 — now**

Master of Engineering, Information and Communication Engineering (master advisor: Chong Fu)
School of Computer Science and Engineering, Northeastern University (China)   **2019.9 — 2022.7**

Bachelor of Engineering, Communication Engineering (GPA: 3.5/5)
School of Computer and Communication Engineering, Northeastern University (China)   **2015.9 — 2019.7**

RESEARCH ACHIEVEMENTS (SELECTED)

My research focuses on **computer vision for scene understanding, including zero-shot, open-vocabulary segmentation based on vision-language model, and multi-modality learning**. The key outcomes of these works are listed as follows:

Journal Publications

1. **Jialei Chen**, Daisuke Deguchi, Chenkai Zhang, Xu Zheng, Hiroshi Murase, **Frozen is Better than Learning: A New Design of Prototype-Based Classifier for Semantic Segmentation**, *Pattern Recognition* (IF 7.5/ JCR Q1)
   https://www.sciencedirect.com/science/article/pii/S0031320324001821

2. **Jialei Chen**, Chong Fu, Haoyu Xie, Xu Zheng, Chiu-Wing Sham, **Uncertainty Teacher with Dense Focal Loss for Semi-Supervised Medical Image Segmentation**, *Computers in Biology and Medicine* (IF 7.0 / JCR Q1)
   https://www.sciencedirect.com/science/article/pii/S001048252200751X

3. Zhenzhen Quan, **Jialei Chen**, Daisuke Deguchi, Jie Sun, Chenkai Zhang, Yujun Li, Hiroshi Murase, **Semantic Matters: A Constrained Approach for Zero-Shot Video Action Recognition**, *Pattern Recognition* (IF 7.5 / JCR Q1)
   https://www.sciencedirect.com/science/article/pii/S0031320325000627

4. Chenkai Zhang, Daisuke Deguchi, **Jialei Chen**, Hiroshi Murase, **Toward Explainable End-to-End Driving Models via Simplified Objectification Constraints**, *IEEE Transactions on Intelligent Transportation Systems* (IF 7.9 / JCR Q1)
   https://ieeexplore.ieee.org/abstract/document/10505932

International Conference

1. **Jialei Chen**, Daisuke Deguchi, Chenkai Zhang, Xu Zheng, and Hiroshi Murase, **Centroid Module for Shaping Feature Space in Semantic Segmentation**, Proceedings of the 3rd Asia Conference on Algorithms, Computing and Machine Learning, 2024.

2. **Jialei Chen**, Daisuke Deguchi, Chenkai Zhang, Zhenzhen Quan, Seigo Ito, Hiroshi Murase, **Multi-group Vision Semantic Centroid for Semantic Segmentation**, International Workshop on Frontiers of Computer Vision, 2025.

3. Zhenzhen Quan, Daisuke Deguchi, **Jialei Chen**, Chenkai Zhang, Yujun Li, Seigo Ito, Hiroshi Murase, **A Cross-Modal Knowledge Distillation Approach for RGB-to-Infrared Video Action Recognition**, International Workshop on Frontiers of Computer Vision, 2025.

4. Chenkai ZHANG, Daisuke DEGUCHI, **Jialei Chen**, Zhenzhen QUAN, Hiroshi MURASE, **CROCODILE: Crop-based Contrastive Discriminative Learning for Enhancing Explainability of End-to-End Driving Models**, Proceedings of the Asian Conference on Computer Vision, 2024.

5. Hiroto Murakami, **Jialei Chen**, Daisuke Deguchi, Takatsugu Hirayama, Yasutomo Kawanishi, Hiroshi Murase, **Pedestrian's Gaze Object Detection in Traffic Scene**, Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 2024.

6. Chenkai Zhang, Daisuke Deguchi, **Jialei Chen**, Hiroshi Murase, **Comprehensive Evaluation of End-to-End Driving Model Explanations for Autonomous Vehicles**, Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 2024.

Domestic Conference

1. <u>Jialei Chen</u>, Daisuke Deguchi, Chenkai Zhang, Xu Zheng, and Hiroshi Murase, **A Study on Fixed Orthogonal Prototype Classifier for Semantic Segmentation**, PRMU, 2024.

2. <u>Jialei Chen</u>, Daisuke Deguchi, Chenkai Zhang, and Hiroshi Murase, **Distill from CLIP by Token Alignment for Inductive Zero-shot Semantic Segmentation**, Proceedings of 27th Meeting on Image Recognition and Understanding (MIRU), 2024.

3. 星谷那月, 出口大輔, 陳嘉雷, 村瀬洋, 伊藤誠悟, 物体属性に関する論理制約を用いた車載カメラ画像からの物体検出, 電子情報通信学会技術研究報告 (MVE), 2025.

4. Chenkai Zhang, Daisuke Deguchi, <u>Jialei Chen</u>, Hiroshi Murase, **A Preliminary Study on Pre-training for Improved Explainability in End-to-End Driving Models**, PRMU 2024.

5. Chenkai Zhang, Daisuke Deguchi, <u>Jialei Chen</u>, Hiroshi Murase, **Enhancing Explainability of End-to-End Autonomous Driving Models through additional fine-tuning**, Proceedings of 27th Meeting on Image Recognition and Understanding (MIRU), 2024.

6. 村上大斗, 陳嘉雷, 出口大輔, 平山高嗣, 川西康友, 村瀬洋, 交通シーン画像からの歩行者の注視対象物推定, 動的画像処理実利用化ワークショップ (DIA2024) 講演論文集.

7. 村上大斗, 陳嘉雷, 出口大輔, 平山高嗣, 川西康友, 村瀬洋, 交通シーンにおける歩行者の注視対象物推定の検討, 第 26 回画像の認識・理解シンポジウム (MIRU2023) 論文集.

Under Review / Preprints

1. <u>Jialei Chen</u>, Daisuke Deguchi, Chenkai Zhang, Xu Zheng, and Hiroshi Murase, **Clip is Also a Good Teacher: A New Learning Framework for Inductive Zero-Shot Semantic Segmentation**, *arXiv preprint*, submitted to *IEEE Transactions on Circuits and Systems for Video Technology*
https://arxiv.org/pdf/2310.02296

2. <u>Jialei Chen</u>, Daisuke Deguchi, Chenkai Zhang, and Hiroshi Murase, **Generalizable Semantic Vision Query Generation for Zero-Shot Panoptic and Semantic Segmentation**, *arXiv preprint*, submitted to *International Journal on Computer Vision*
https://arxiv.org/pdf/2402.13697

3. **Training-Free Open-Vocabulary Semantic Segmentation with Affinity Pyramid Refinement** (First Author)
Submitted to *ICCV 2025*.

4. **BiXFormer: A Robust Framework for Maximizing Modality Effectiveness in Multi-Modal Semantic Segmentation** (First Author)
Submitted to *ICCV 2025*.

5. <u>Jialei Chen</u>, Dongyue Li, Chong Yi, Ito Seigo, Daisuke Deguchi, Hiroshi Murase, **Unseen-aware Hungarian Matching for Inductive Zero-shot Semantic Segmentation**, IEEE International Conference on Systems, Man, and Cybernetics, Under Review.

6. <u>Jialei Chen</u>, Dongyue Li, Chong Yi, Ito Seigo, Daisuke Deguchi, Hiroshi Murase, **Semantic-driven Distillation for Transductive Zero-shot Semantic Segmentation**, IEEE International Conference on Intelligent Transportation Systems 2025, Under Review.

RESEARCH CONTENTS

Zero-shot Segmentation, Open-vocabulary Segmentation based on Vision-language Model, Multi-modality Learning, etc.

SKILLS AND EXPERTISE

- **Ability to Plan and Execute Projects**

  – Literature investigation, experiment implementation, results analysis.

- **Strong Writing and Presentation Skills**

  – **Paper writings:**
  Published several first-authored and co-authored papers in top-tier journals and conferences.
  Submitted six times to top-tier conferences since 2023.
  – **Presentations:**
  Oral and poster presentations at international and domestic conferences.

- **Expertise in Computer Vision Research**

  – Specialized in **semantic segmentation**, including conventional, zero-shot, open-vocabulary, and multi-modality approaches.
  – Experience with **foundation models**, including vision-language and vision-only architectures.

- **Hands-on Programming Experience**
  - Conducting experiments using **PyTorch**.
  - Performing model benchmarking with **MMsegmentation** and **MMdetection** frameworks.
  - Managing code versioning and collaboration via **GitHub**.

Connection between my research and the research topic in CyberAgent

I am applying for the research topic **"Multimodal Foundation Model for Human Action Understanding"** offered by CyberAgent, which strongly alignes with my own research on **human-centered environmental understanding**. I understand this topic as focusing on two key goals: (1) interpreting user behavior through their interactions with the real world, and (2) leveraging multi-modal inputs, such as visual, textual, and auditory data, to support robust understanding across different scenarios.

To solve (1), I am working on **semantic segmentation**, which enables pixel-wise understanding of both users and the objects around them, providing a crucial foundation for modeling human-object interactions. For example, in a retail environment, recognizing someone reaching for a shelf or holding a shopping basket can help infer their intention.

To solve (2), my research goes beyond fixed object categories and explores **text-driven zero-shot semantic segmentation** using vision-language models like CLIP. By aligning pixel-level features with natural language, the model can recognize rare but behaviorally important objects such as wheelchairs or strollers, without relying on exhaustive annotations. I am also extending segmentation to **non-traditional modalities** like LiDAR and event cameras for robust scene understanding.

In summary, CyberAgent's research aims to understand human actions using multimodal inputs, a goal that closely aligns with my own work on building robust and generalizable perception systems. With experience in **semantic segmentation, zero-shot learning, and multimodal modeling**, I am confident in my ability to contribute to this research topic.

Academic Services

Reviewers for CVPR, IJCNN, IEEE Signal Processing Letters

Awards

1. First class academic scholarship, Northeastern University, 2019.

2. Second class academic scholarship, Northeastern University, 2020.

3. 陳嘉雷, Towards Zero-Shot Semantic Segmentation based on Vision-Language Model, 教員が選ぶ優秀発表賞, 2024.

4. 村上大斗, 陳嘉雷, 出口大輔, 平山高嗣, 川西康友, 村瀬洋, 交通シーン画像からの歩行者の注視対象物推定, 研究奨励賞, 2024.

5. Chenkai Zhang, Daisuke Deguchi, **Jialei Chen**, Hiroshi Murase, A Preliminary Study on Pre-training for Improved Explainability in End-to-End Driving Models, PRMU 月間ベストプレゼンテーション賞, 2024.

Language

Japanese: Conversational, English: Fluent, Chinese: Native.