Fangyi Chen

Research Scientist @ Intelligent Creation, ByteDance

Tel: (412)-417-5016 Email: fangyichen5@gmail.com

Research Interests

- Artificial intelligence, Deep learning
- Object detection, Multimodal large language model, Open-world scene understanding

Education

- Ph.D. in Electrical and Computer Engineering — 2020-2025
  Carnegie Mellon University — Pittsburgh, USA
  Advisor: Marios Savvides
- M.S. in Electrical Engineering — 2017-2018
  University of Pittsburgh — Pittsburgh, USA
- B.E. in Electrical Engineering and Its Automation — 2013-2017
  North China Electric Power University — Beijing, China

Publications (Google Scholar)

[1] Hao Chen, Yujin Han, **Fangyi Chen**, Xiang Li, Yidong Wang, Jindong Wang, Ze Wang, Zicheng Liu, Difan Zou, Bhiksha Raj. Masked autoencoders are effective tokenizers for diffusion models. *International Conference on Machine Learning (ICML), 2025.*

[2] Hao Chen, Ze Wang, Xiang Li, Ximeng Sun, **Fangyi Chen**, Jiang Liu, Jindong Wang, Bhiksha Raj, Zicheng Liu, Emad Barsoum. Softvq-vae: Efficient 1-dimensional continuous tokenizer. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2025.*

[3] **Fangyi Chen***, Han Zhang*, Zhantao Yang, Hao Chen, Kai Hu, Marios Savvides. RTGen: Generating Region-Text Pairs for Open-Vocabulary Object Detection. *(preprint)*

[4] Yu-Kai Huang, Yutong Zheng, Yen-Shuo Su, Anudeepsekhar Bolimera, Han Zhang, **Fangyi Chen**, Marios Savvides. A Reference-Based 3D Semantic-Aware Framework for Accurate Local Facial Attribute Editing. *IEEE International Joint Conference on Biometrics (IJCB), 2024.*

[5] **Fangyi Chen**, Han Zhang, Kai Hu, Yu-kai Huang, Chenchen Zhu, Marios Savvides. Enhanced Training of Query-Based Object Detection via Selective Query Recollection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2023.*

[6] **Fangyi Chen**, Han Zhang, Zaiwang Li, Jiachen Dou, Shentong Mo, Hao Chen, Yongxin Zhang, Uzair Ahmed, Chenchen Zhu, Marios Savvides. Unitail: Detecting, Reading, and Matching in Retail Scene. *European Conference on Computer Vision (ECCV), 2022.*

[7] Chenchen Zhu, **Fangyi Chen**, Uzair Ahmed, Zhiqiang Shen, Marios Savvides. Semantic Relation Reasoning for Shot-Stable Few-Shot Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021.*

[8] Chenchen Zhu, **Fangyi Chen**, Zhiqiang Shen, Marios Savvides. Soft Anchor-Point Object Detection. *European Conference on Computer Vision (ECCV), 2020.*

[9] **Fangyi Chen**, Chenchen Zhu, Zhiqiang Shen, Han Zhang, Marios Savvides. NCMS: Towards Accurate Anchor Free Object Detection through l2 Norm Calibration and Multi-Feature Selection. *Computer Vision and Image Understanding (CVIU), 2020 Jul 27:103050.*

[10] Han Zhang, **Fangyi Chen**, Zhiqiang Shen, Qiqi Hao, Chenchen Zhu, Marios Savvides. Solving Missing-Annotation Object Detection With Background Recalibration Loss. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2020.*

[11] **Fangyi Chen**, Chenchen Zhu, Marios Savvides. A Novel Collaborate Control Strategy for Enhanced Training of Vehicle Recognition. *The IEEE 90th Vehicular Technology Conference (VTC), 2019.*

[12] Ker-Jiun Wang, Kaiwen You, **Fangyi Chen**, Prakash Thakur, Michael Urich, Soumya Vhasure, and Zhi-Hong Mao. Development of Seamless Telepresence Robot Control Methods to Interact with The Environment Using Physiological Signals. *The 13th ACM International Conference on Human-Robot Interaction (HRI), 2018.*

[13] Ker-Jiun Wang, Anna Zhang, Kaiwen You, **Fangyi Chen**, Quanbo Liu, Yu Liu, Zaiwang Li, Hsiao-Wei Tung, and Zhi-Hong Mao. Ergonomic and Human Centered Design of Wearable Gaming Controller Using Eye Movements and Facial Expressions. *The IEEE International Conference on Consumer Electronics (ICCE), 2018.*

[14] Ker-Jiun Wang, Kaiwen You, **Fangyi Chen**, Zihang Huang, and Zhi-Hong Mao. Human-Machine Interface Using Eye Saccade and Facial Expression Physiological Signals to Improve the Maneuverability of Wearable Robots. *The International Symposium on Wearable & Rehabilitation Robotics (WeRob 2017), 2017.*

Patents

[1] US12189714B2	System and method for improved few-shot object detection using dynamic semantic network
[2] US20250182450A1	System and method for weapon detection with pose estimation
[3] US12266156B2	System and method for solving missing annotation object detection
[4] US20250005881A1	System and method for assigning complex concave polygons as bounding boxes
[5] US12131497B2	Fast object search based on the cocktail party effect
[6] US20240355085A1	System and method for matching products and determining spreads and plugs
[7] US11915463B2	System and method for the automatic enrollment of object images into a gallery
[8] WO2020210825A1	System and method for detecting products and product labels
[9] WO2022211995A1	System and method for using non-axis aligned bounding boxes for retail detection
[10] WO2022169622A1	Soft anchor point object detection
[11] WO2022109295A1	System and method for detecting and classifying abnormal cells
[12] US11954175	Feature pyramids for object detection
[13] US2022058432A1	Few-shot object detection using semantic relation reasoning

Professional Experience

Research Scientist in Generative AI @ ByteDance, Bellevue 07/2025 - present

Journal Reviewer:

- IEEE Transactions on Image Processing (since 2022)
- ELSEVIER Pattern Recognition (since 2020)
- IEEE Transactions on Geoscience and Remote Sensing (since 2020)
- ELSEVIER NeuralComputing (since 2022)
- Springer Visual Computer (since 2022)
- Connected Science (since 2022)
- Multimedia Systems (since 2020)
- Springer International Journal of Computer Vision (since 2025)

Conference Reviewer:

- The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2023, 2024, 2025
- The Annual Conference on Neural Information Processing Systems (NeurIPS) 2024, 2025
- International Conference on Computer Vision (ICCV) 2023, 2025
- European Conference on Computer Vision (ECCV) 2024
- The International Conference on Learning Representations (ICLR) 2023, 2024, 2025
- International Conference on Machine Learning (ICML) 2023, 2024, 2025

Research Intern (Intelligent Creation-Vision and Graphics) @ ByteDance, Bellevue 05/2024-08/2024

- **Multimodal LLM**: Conducting research on multi-modal large language model, aiming to design and train AI system for in-depth image editing, decomposition, and synthesis.

Research Assistant (2020-present) & Research Associate III (2019) @ Cylab, CMU 02/2019-05/2025

- **U.S. Department of Defence, Project Maven:** Object detection training strategy with insufficient and imperfect data.
- **AI for Retail:** Real-time system and robot for store management and automatic checkout. Lead a team for large-scale dataset collection and annotation. Develop RetailDet, a quadrilateral product detector that achieves top performance on three retail datasets, and a textually enhanced product matching algorithm that operates in a one-shot manner. The system is deployed in 350 Walmart stores. Research results are turned into patents.
- **Anchor-free detection system and robots:** A progressive design of anchor-free object detection aims to address the inherently heuristic feature selection of anchor-based detectors. They are developed with novel ground-truth assignment strategies on feature map and across feature pyramid via multi-level feature selection, norm calibration, and soft-weighted training losses. Research results are turned into patents.

Selected Project @ ECE, UPitt 09/2017-03/2018

- **EXGbuds:** The development of a wearable device and machine learning algorithm to measure eye movements and facial expressions to generate useful commands via non-invasive biosensors. The research result is adopted by EXGwear Inc.

Awards

- Carnegie Institute of Technology Dean's Fellowship, Carnegie Mellon University 2020
- Best Hardware Hack, PITT-CTSI GitHub Major League Hacking 2018
- National Third Prize, Chinese National College Students Competition on Energy Economics 2015
- University Merit Student, NCEPU 2014