

# Qiyu Wu

Tokyo, Japan | [wuqiyu576@gmail.com](mailto:wuqiyu576@gmail.com) | Personal Website: [qiyuw.github.io](https://qiyuw.github.io) | LinkedIn: [qiyuw](#)

Keywords: NLP; Multilingual; Multimodal; LLMs.

## Summary

Tech lead of Multimodal NLP team at Creative AI Lab, Sony. Ph.D. specializing in NLP, focusing on **Multimodal Understanding and Representation Learning**. These works have been presented at conferences such as **ACL**, **EMNLP**, **ICLR**, **NAACL**, **ICML**, **TMLR**, **VLDB**, **EACL**, **AAAI**, etc. I also served as area chair and reviewers in above top conferences and organized workshops at NeurIPS. I received JSPS DC fellowship during my Ph.D.

## Research Experiences

<b>Research Scientist and Tech Lead at Sony</b>	2025.10 – Present
<ul style="list-style-type: none"><li>Leading Multimodal NLP team, Creative AI Lab.</li><li>Multi-modal pre-training for understanding and retrieval [M1, C12, C14, C16].</li></ul>	
<b>Ph.D. at UTokyo, Tsuruoka Lab</b>	2021.10 – 2024.09
<ul style="list-style-type: none"><li>Word-level preference for LLM-based machine translation [C11]; Word alignment pre-training [C5, C15];</li><li>Cross-lingual text embeddings [C8, C7].</li></ul>	
<b>Visiting Student at UC San Diego</b>	2023.01 – 2023.02
<ul style="list-style-type: none"><li>Generalized diffusion model [C9].</li></ul>	
<b>Research Intern at Microsoft</b>	2020.04 – 2022.02
<ul style="list-style-type: none"><li>Multiple positive examples in sentence embeddings with contrastive learning [C4], work done at STCA.</li><li>Acceleration of language pre-training via better rare word representation [C3], work done at MSR Asia.</li></ul>	
<b>R&amp;D Intern at Baidu Research</b>	2019.10 – 2020.02
<ul style="list-style-type: none"><li>Multi-task transportation prediction [C2]; Route mode recommendation in Baidu Maps.</li></ul>	

## Education

<b>The University of Tokyo</b> , Ph.D. Degree in Information Science & Technology.	2021.10 – 2024.09
<ul style="list-style-type: none"><li>Ph.D. at Tsuruoka Lab; JSPS DC2 Fellowship; JST SPRING Fellowship.</li></ul>	
<b>Peking University</b> , M.Sc. Degree in Computer Software & Theory.	2018.09 – 2021.07
<b>Sichuan University</b> , B.Eng. Degree in Computer Science & Technology.	2014.09 – 2018.07

## Advising and Services

<b>Advised Students</b>	
<ul style="list-style-type: none"><li>Jongsuk Kim (2025), Intern at Sony, Ph.D. student from KAIST, published [M1].</li><li>Zhongtao Miao (2023 – 2024), Ph.D. student at The University of Tokyo, published [C8, C15].</li><li>Kaiyan Zhao (2022 – 2024), Ph.D. student at The University of Tokyo, published [C7, C13].</li><li>Yutao Xie (2020 – 2021), Ph.D. student at Peking University, published [J1].</li></ul>	
<b>Area Chair of ACL Rolling Review.</b>	2025 – Present
<b>Reviewer of ACL Rolling Review, EMNLP, ICLR, NeurIPS, ICML, etc.</b>	2021 – Present
<b>Organizer of “NeurIPS 2025 GenProCC Workshop”.</b>	2025
<b>DC2 Fellowship of JSPS.</b>	2024
<b>Teaching Assistant of “Social Implications of Computing” at Peking University.</b>	2019

## Publications

---

\* indicates equal contribution; † indicates students mentored by me.

[M1] Jongsuk Kim<sup>†</sup>, Qiyu Wu, Zhuoyuan Mao, Hiromi Wakaki, Junmo Kim, Yuki Mitsufuji. “MLLMCLIP: Distilling Multimodal Language Models for Robust Vision–Language Understanding”, *manuscript*.

[C16] Zhuoyuan Mao, Mengjie Zhao, Qiyu Wu, Hiromi Wakaki, Yuki Mitsufuji. “DeepResonance: Enhancing Multimodal Music Understanding via Music-centric Multi-way Instruction Tuning”, (*EMNLP 2025 Main*).

[C15] Zhongtao Miao<sup>†</sup>, Qiyu Wu, Masaaki Nagata, and Yoshimasa Tsuruoka. “Improving Word Alignment Using Semi-Supervised Learning”, (*ACL 2025 Findings*).

[C14] Zhuoyuan Mao, Mengjie Zhao, Qiyu Wu, Zhi Zhong, Wei-Hsiang Liao, Hiromi Wakaki, Yuki Mitsufuji. “Cross-Modal Learning for Music-to-Music-Video Description Generation”, in Proceedings of the 10th Workshop on Representation Learning for NLP (*RepL4NLP-2025*).

[C13] Kaiyan Zhao<sup>†</sup>, Qiyu Wu, Zhongtao Miao, Yoshimasa Tsuruoka. “Prompt Tuning Can Simply Adapt Large Language Models to Text Encoders”, in Proceedings of the 10th Workshop on Representation Learning for NLP (*RepL4NLP-2025*).

[J2] Yuting Tang, Xin-Qiang Cai, Yao-Xiang Ding, Qiyu Wu, Guoqing Liu, Masashi Sugiyama. “Reinforcement Learning from Bagged Reward”, in Transactions on Machine Learning Research (*TMLR*).

[C12] Qiyu Wu, Mengjie Zhao, Yutong He, Lang Huang, Junya Ono, Hiromi Wakaki, Yuki Mitsufuji. “Towards Reporting Bias in Visual-language Datasets: Bi-modal Data Augmentation by Decoupling Object-attribute Association”, (*ICCV 2025 Workshop Multimodal Representation and Retrieval*).

[C11] Qiyu Wu, Masaaki Nagata, Zhongtao Miao, Yoshimasa Tsuruoka. “Word Alignment as Preference for Machine Translation”, (*EMNLP 2024, main conference long paper; Oral*).

[C10] Qiyu Wu, Zilong Wu, Yoshimasa Tsuruoka. “SGA: Scene Graph Alignment for Evaluation of Text-to-Image Generation”, (*KDD 2024 Workshop HI-AI*).

[C9] Guangyi Liu, Yu Wang, Zeyu Feng, Qiyu Wu, Liping Tang, Yuan Gao, Zhen Li, Shuguang Cui, Julian McAuley, Eric P. Xing, Zichao Yang, Zhiting Hu. “Unified Generation, Reconstruction, and Representation: Generalized Diffusion with Adaptive Latent Encoding-Decoding”, in the Forty-first International Conference on Machine Learning. (*ICML 2024*).

[C8] Zhongtao Miao<sup>†</sup>, Qiyu Wu, Kaiyan Zhao, Zilong Wu, Yoshimasa Tsuruoka. “Enhancing Cross-lingual Sentence Embedding for Low-resource Languages with Word Alignment”, in the 2022 Conference of the North American Chapter of the Association for Computational Linguistics (*NAACL 2024, findings*).

[C7] Kaiyan Zhao<sup>†\*</sup>, Qiyu Wu<sup>\*</sup>, Xin-Qiang Cai, Yoshimasa Tsuruoka. “Leveraging Multi-lingual Positive Instances in Contrastive Learning to Improve Sentence Embedding”, in the 18th Conference of the European Chapter of the Association for Computational Linguistics (*EACL 2024, main conference long paper*).

[C6] Shiwen Wu, Qiyu Wu, Honghua Dong, Wen Hua, Xiaofang Zhou. “Blocker and Matcher Can Mutually Benefit: A Co-Learning Framework for Low-Resource Entity Resolution.”, *Proc. VLDB Endow.* (2024). Will appear at Very Large Databases Conf. (*VLDB 2024, research track paper*).

[C5] Qiyu Wu, Masaaki Nagata, Yoshimasa Tsuruoka. “WSPAlign: Word Alignment Pre-training via Large-Scale Weakly Supervised Span Prediction”, in the 61st Annual Meeting of the Association for Computational Linguistics (*ACL 2023, main conference long paper*).

[C4] Qiyu Wu, Chongyang Tao, Tao Shen, Can Xu, Xiubo Geng, Daxin Jiang. “PCL: Peer-Contrastive Learning with Diverse Augmentations for Sentence Embeddings”, in *Conference on Empirical Methods in Natural Language Processing (EMNLP 2022, main conference long paper)*.

[J1] Yutao Xie<sup>†</sup>, Qiyu Wu, Wei Chen and Tengjiao Wang. “Stable Contrastive Learning for Self-Supervised Sentence Embeddings With Pseudo-Siamese Mutual Learning”, in *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, vol. 30, pp. 3046-3059, 2022.

[C3] Qiyu Wu\*, Chen Xing\*, Yatao Li, Guolin Ke, Di He, Tie-Yan Liu. “Taking Notes on the Fly Helps Language Pre-Training”, in *the Ninth International Conference on Learning Representations (ICLR 2021)*.

[C2] Hao Liu\*, Qiyu Wu\*, Fuzhen Zhuang, Xinjiang Lu, Dejing Dou, Hui Xiong. “Community-Aware Multi-Task Transportation Demand Prediction”, in *Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI 2021)*.

[C1] Dechuan Yang\*, Qiyu Wu\*, Wei Chen, Tengjiao Wang, Zhen Qiu, Di Liu. “Stance Detection with Stance-wise Convolution Network”, in the *2022 CCF International Conference on Natural Language Processing and Chinese Computing*.

[J]: Journal paper; [C]: Conference paper; [M]: Manuscript; [N]: Non-refereed.