Yun-Cheng (Joe) Wang

joewang622@gmail.com
(213) 379-0669
555 San Antonio Rd. Apt 320
Mountain View, CA, 94040

Career Objectives

Research and education in various topics in artificial intelligence (AI) and machine learning (ML), with a focus on bridging fundamental innovations and real-world engineering applications. I am passionate about transforming foundational ideas into scalable and intelligent systems.

Education

University of Southern California – Los Angeles, CA — Jan 2021 - Dec 2023
Ph.D. in Electrical Engineering, Minor area in Computer Science

- Dissertation: *Green Knowledge Graph Completion and Scalable Generative Content Delivery*

University of Southern California – Los Angeles, CA — Aug 2018 - Dec 2019
M.S. in Electrical Engineering

National Taiwan University – Taipei, Taiwan — Sep 2014 - Jun 2018
B.S. in Electrical Engineering

Research Area

- **Knowledge Representation and Reasoning**
  - Reasoning over knowledge graphs
  - Agent-based reasoning with large language models
  - Multi-modal information extraction and retrieval
- **Efficient ML Methods and Scalable AI Systems**
  - Representation learning for high-dimensional data
  - Time- and memory-efficient ML pipelines for real-world applications
  - Edge–cloud collaboration for privacy-preserving AI

Employment

Senior Machine Learning Engineer, Apple, Inc. – Cupertino, CA — Jan 2025 - Present
Siri Natural Language Data Science

Research Scientist, Yahoo, Inc. – Mountain View, CA — Jan 2024 – Jan 2025
Search and Recommendation Science

Publications

Conference Papers

[1] Jintang Xue, **Yun-Cheng Wang**, Chengwei Wei, C.-C. Jay Kuo, "Efficient Feature Selection for Word Embedding Dimension Reduction," *Asia Pacific Signal and Information Processing Association Annual Summit and Conference (ASC)*, 2024.

[2] Zhanxuan Mei, **Yun-Cheng Wang**, C.-C. Jay Kuo, "GSBIQA: Green Saliency-guided Blind Image Quality Assessment Method," *Asia Pacific Signal and Information Processing Association Annual Summit and Conference (ASC)*, 2024.

[3] Tsung-Shan Yang, **Yun-Cheng Wang**, Chengwei Wei, Suya You, C.-C. Jay Kuo, "GMA: Green Multi-Modal Alignment for Image-Text Retrieval," *Asia Pacific Signal and Information Processing Association Annual Sum-*

mit and Conference (ASC), 2024.

[4] Tsung-Shan Yang, **Yun-Cheng Wang**, Chengwei Wei, C.-C. Jay Kuo, "GHOI: A Green Human-Object-Interaction Detector," *IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2024.

[5] **Yun-Cheng Wang**, Xiou Ge, Bin Wang, C.-C. Jay Kuo, "AsyncET: Asynchronous Representation Learning for Knowledge Graph Entity Typing," *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.

[6] **Yun-Cheng Wang**, Xiou Ge, Bin Wang, C.-C. Jay Kuo, "GreenKGC: A Lightweight Knowledge Graph Completion Method," *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.

[7] Xiou Ge, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "Compounding Geometric Operations for Knowledge Graph Completion," *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.

[8] Zhanxuan Mei, **Yun-Cheng Wang**, Xingze He, C.-C. Jay Kuo, "GreenBIQA: A Lightweight Blind Image Quality Assessment Method," *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2022. **(Top 10% Paper Award)**

[9] Bin Wang, Fenxiao Chen, **Yun-Cheng Wang**, C.-C. Jay Kuo, "Efficient Sentence Embedding via Semantic Subspace Analysis," *IEEE International Conference on Pattern Recognition (ICPR)*, 2020.

Journal Papers

[1] Tsung-Shan Yang, **Yun-Cheng Wang**, Chengwei Wei, Suya You, C.-C. Jay Kuo, "Efficient Human-Object-Interaction (EHOI) Detection via Interaction Label Coding and Conditional Decision," *Computer Vision and Image Understanding*, 2025.

[2] Jintang Xue, **Yun-Cheng Wang**, Chengwei Wei, Xiaofeng Liu, Jonghye Woo, C.-C. Jay Kuo, "Word Embedding Dimension Reduction via Weakly-Supervised Feature Selection," *APSIPA Transactions on Signal and Information Processing*, 2024.

[3] Zhanxuan Mei, **Yun-Cheng Wang**, C.-C. Jay Kuo, "Blind Video Quality Assessment at the Edge," *IEEE Transactions on Multimedia*, 2024.

[4] Zhanxuan Mei, **Yun-Cheng Wang**, C.-C. Jay Kuo, "GreenSaliency: A Lightweight and Efficient Image Saliency Detection Method," *APSIPA Transactions on Signal and Information Processing*, 2024.

[5] Xiou Ge, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "Knowledge Graph Embedding: An Overview," *APSIPA Transactions on Signal and Information Processing*, 2024.

[6] Jintang Xue, **Yun-Cheng Wang**, Chengwei Wei, Xiaofeng Liu, Jonghye Woo, C.-C. Jay Kuo, "Bias and Fairness in Chatbots: An Overview," *APSIPA Transactions on Signal and Information Processing*, 2024.

[7] Xiou Ge, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "Knowledge Graph Embedding with 3D Compound Geometric Transformations," *APSIPA Transactions on Signal and Information Processing*, 2024.

[8] Zhanxuan Mei*, **Yun-Cheng Wang***, Xingze He, Yong Yan, C.-C. Jay Kuo, "Lightweight High-Performance Blind Image Quality Assessment," *APSIPA Transactions on Signal and Information Processing*, 2024.

[9] **Yun-Cheng Wang**, Jintang Xue, Chengwei Wei, C.-C. Jay Kuo, "An Overview on Generative AI at Scale with Edge-Cloud Computing," *IEEE Open Journal of the Communications Society*, 2023.

[10] Chengwei Wei, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "An Overview on Language Models: Recent Developments and Outlook," *APSIPA Transactions on Signal and Information Processing*, 2023.

[11] Xiou Ge, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "TypeEA: Type-Associated Embedding for Knowledge Graph Entity Alignment," *APSIPA Transactions on Signal and Information Processing*, 2023.

[12] **Yun-Cheng Wang**, Xiou Ge, Bin Wang, C.-C. Jay Kuo, "KGBoost: A Classification-Based Knowledge Base Completion Method with Negative Sampling," *Pattern Recognition Letters*, 2022.

[13] Xiou Ge, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "CORE: A Knowledge Graph Entity Type Prediction Method via Complex Space Regression and Embedding," *Pattern Recognition Letters*, 2022.

[14] Fenxiao Chen, **Yun-Cheng Wang**, Bin Wang, C.-C. Jay Kuo, "Graph Representation Learning: A Survey,"

APSIPA Transactions on Signal and Information Processing, 2020. **(2024 Sadaoki Furui Prize Paper Award)**

[15] Bin Wang, Angela Wang, Fenxiao Chen, **Yun-Cheng Wang**, C.-C. Jay Kuo, "Evaluating Word Embedding Models: Methods and Experimental Results," *APSIPA Transactions on Signal and Information Processing*, 2019. **(2022 Sadaoki Furui Prize Paper Award)**

Invited Talks

- "Scalable Generative Content Delivery", Research Center for Information Technology Innovation (CITI) at Academia Sinica, Jan 2024.
- "Green Learning with Knowledge Graphs", National Taiwan University (NTU), Dec 2023.
- "Knowledge Graph: Foundations and Applications", National Taiwan Normal University (NTNU), Sep 2020.

Awards

- USC Viterbi Graduate Fellowship/Research Assistantship/Teaching Assistantship — Jan 2021 - Dec 2023
- Sadaoki Furui Prize Paper Award (Pioneering Contributions in Speech Processing) — APSIPA ASC 2022, 2024
- Top 10% Paper Award — IEEE MMSP 2022
- USC EE MS honors program — Aug 2019 - Dec 2019

Academic Services

Conference Reviewer / Program Committee

- Conference on Language Modeling (CoLM): 2025
- International Conference on Knowledge Discovery and Data Mining (KDD): 2024, 2025
- ACL Rolling Review (ARR): Feb 2024, Apr 2024, Jun 2024, Aug 2024
- Empirical Methods in Natural Language Processing (EMNLP): 2023
- Asian Conference on Machine Learning (ACML): 2024
- European Conference on Machine Learning (ECML): 2022
- International Conference on Computational Linguistics (COLING): 2024, 2025
- International Conference on Acoustics, Speech, and Signal Processing (ICASSP): 2025

Journal Reviewer

- IEEE/ACM Transactions on Audio, Speech and Language Processing (T-ASL)
- IEEE Internet of Things Magazine (IoTM)
- IEEE Journal of Biomedical and Health Informatics (JBHI)
- IEEE Transactions on Artificial Intelligence (TAI)
- IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)
- IEEE Consumer Electronics Magazine (CEM)
- APSIPA Transactions on Signal and Information Processing

Professional Affiliations

- Member, ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD)
- Member, Association for Computational Linguistics (ACL)
- Member, IEEE Signal Processing Society

Research Projects

Improving Interpretability of Language Models — Aug 2023 - Present

- Integrated extracted knowledge from data with lightweight language models for language generation

- Developed efficient and reliable generative models for high-stakes domains such as biomedical applications.
- Investigated local n-gram and constituency-dependent n-gram models to simplify the self-attention mechanism.

Multi-modal Retrieval and Alignment, *Sponsored by Army Research Lab (ARL)* Aug 2023 - Present

- Leveraged the embedding space to connect different modalities for multi-modal reasoning.
- Extracted human-object interactions using spatial and latent features with hierarchical classifiers.
- Devised an alignment module in a joint embedding space for text-to-image and image-to-text retrieval.

Entity Summarization with Knowledge Graphs, *Intern work with Yahoo Inc.* Jun 2023 - Aug 2023

- Generated knowledge-grounded entity descriptions in knowledge graphs.
- Innovated a fact-ranking system to retrieve relevant relationships of entities.
- Constructed a large-scale synthetic data-to-text dataset for model training and evaluation.

Efficient Reasoning on KGs using Lightweight Models Jan 2021 - Dec 2023

- This project aimed at predicting missing information, including entity types and relations, in KGs using lightweight models.
- Adopted feature pruning to achieve parameter efficiency and SOTA performance in low dimensions.
- Proposed novel modeling of entity types to improve expressiveness while retaining scalability to large KGs.
- Innovated an asynchronous KGE learning framework and achieved SOTA performance for embedding-based models in both link prediction and entity type prediction tasks.
- Reduced the number of inference Floating Point Operations (FLOPs) by over 100$\times$, and the number of parameters by over 15$\times$.

Scalable Generative AI Services under Edge-Cloud Computing Jan 2023 - Oct 2023

- Analyzed the memory, computation, and network requirements to deploy GenAI services, e.g., ChatGPT, across different scales.
- Estimated the latency for GenAI services under different communication frameworks.
- Identified considerations when designing GenAI systems with better efficiency, computation offloading, and privacy.

Blind Image and Video Quality Assessment, *Sponsored by Meta Inc.* Aug 2021 - Dec 2022

- Used ML models to predict users' perceptual experience on images and videos without the need for reference data, enhancing real-time user experience across various media formats.
- Developed a lightweight, real-time inference pipeline optimized for deployment on user devices, ensuring seamless integration with minimal resource usage.
- Achieved SOTA performance on synthetic datasets and competitive results on authentic datasets, while reducing model size by 54x, significantly improving efficiency and scalability.

Teaching Experiences

Guest Lecturer - University of Southern California Fall 2022, Fall 2023

- Course: Multimedia Compression

Teaching Assistant - University of Southern California Spring 2023 - Fall 2023

- Course: Stochastic Process for Financial Engineering
- Facilitated course delivery and provided guidance through discussions and office hours

Course Mentor - University of Southern California Fall 2019

- Course: Probability for Electrical and Computer Engineers
- Supported course design and hosted office hours