Chenguang Wang

Homepage | Google Schoalr | LinkedIn | Email: chenguang.wang@stonybrook.edu

Education

John Hopkins University, Graduate Visiting Scholar — Aug 2024 – May 2025
Stony Brook University, Ph.D. Candidate in Civil Engineering | GPA: 3.92/4.0 — Jan 2023 – Dec 2025
Stevens Institute of Technology, M.S. Computer Science | GPA: 3.88/4.0 — Aug 2021 – Jan 2023
Xi'an Jiaotong University, B.S. Computer Science — Aug 2016 – June 2020

Research Experience

Stony Brook University — Jan 2023 – Dec 2025
- Supervisor: Prof. Susu Xu & Prof. Ruwen Qin
- Focus: LLM/VLM Post-training and applications in resilience and disaster response.
- More specifically, (i) **LLM applications in Disaster Response**: *CrisisNLP*, *FLARE*; (ii) **LLM Post-training**: *Mosaic-IT*, *RuleR*; (iii) **MLLMs**: *TRIG*, *CaughtCheating*; (iv) **Retrieval-Augmented Generation (RAG)**: *FLARE*; (v) **Agent and RL**: *StreetDesinger*, *SILIC*;

Johns Hopkins University — Aug 2024 – May 2025
- Supervisor: Prof. Susu Xu
- Focus: LLM/VLM applications in resilience and disaster response.

Stevens Institute of Technology — Aug 2021 – Jan 2023
- Supervisor: Prof. Xiaojiang Du
- Focus: Computer Vision, Robotics.

Xi'an Jiaotong University — Jan 2020 – Dec 2020
- Supervisor: Prof. Zhongmin Cai
- Focus: Human-Computer Interaction.

Work Experience

Teaching Assistant, Stony Brook University — Sep 2023 – Dec 2023
- CIV 101 Fall 2023: *Introduction to Civil Engineering* | **Instructor**: Prof. Marija Krstić
- CIV 305 Fall 2023: *Transportation Sytem Analysis*| **Instructor**: Prof. Anil Yazici

Teaching Assistant, Stony Brook University — Jan 2024 – May 2024
- CIV 355 Spring 2024: *Data Analytics for Civil Engineering Systems* | **Instructor**: Prof. Ruwen Qin

Machine Learning Engineering Intern, Zuoyebang Education Technology — Dec 2020 – May 2021
- User next-day activity prediction via ML modeling

Publications

[1] **Wang, C.**, Chen, R., Sun, Y., Zhao, X., & Xu, S. (2025). *From Perceptions to Decisions: Wildfire Evacuation Decision Prediction with Behavioral Theory-informed LLMs*. *ACL 2025*.
[2] **Wang, C.**, Li, M., Chen, H., Nguyen, D., Li, D., & Zhou, T. (2024). *RuleR: Improving LLM Controllability by Rule-based Data Recycling*. *NAACL 2025*.
[3] Li, M., Chen, P., **Wang, C.**, Zhao, H., Liang, Y., Hou, Y., Liu, F., & Zhou, T. (2024). *Mosaic IT: Enhancing instruction tuning with data mosaics*. *ACL 2025*.
[4] **Wang, C.**, Engler, D., Li, X., Hou, J., Wald, D.J., Jaiswal, K., & Xu, S. (2024). *Near-real-time earthquake-induced fatality estimation using crowdsourced data and large-language models*. *International Journal of Disaster Risk Reduction*.
[5] **Wang, C.**, Liu, Y., Zhang, X., Li, X., Paramygin, V., Sheng, P., Zhao, X., & Xu, S. (2024). *Scalable and rapid building damage detection after Hurricane Ian using causal Bayesian networks and InSAR imagery*. *International Journal of Disaster Risk Reduction*.
[6] **Wang, C.**, Liu, Y., Zhang, X., Li, X., Paramygin, V., Subgranon, A., Sheng, P., Zhao, X., & Xu, S. (2023). *Causality-informed rapid post-hurricane building damage detection in large scale from InSAR imagery*. *ACM SIGSPATIAL 2023 EM-GIS Workshop*.
[7] **Wang, C.**, Li, M., Liang, Y., Wang, X., Zhou, Y., Wu, X., Zhang, Y., Zhang, R., & Zhou, T. (2025). *CaughtCheating: Is your MLLM a good cheating detective? Exploring the boundary of visual perception and reasoning*. *NeurIPS 2025 LAW Workshop*.
[8] Sun, Y., Xu, S., **Wang, C.**, & Zhao, X. (2025). *Where You Go is Who You Are: Behavioral Theory-Guided LLMs for Inverse Reinforcement Learning*. *2026 TRB Annual Meeting*.
[9] **Wang, C.**, Yan, X., Dai, Y., Wang, Z., & Xu, S. (2025). *From image generation to infrastructure design: A multi-agent pipeline for street design*. *NeurIPS 2025 Urban AI Workshop*.
[10] Chen, J., Li, M., Kil, J., **Wang, C.**, Yu, T., Rossi, R., Zhou, T., Chen, C., & Zhang, R. (2025). *VisR-Bench: An empirical study on visual retrieval-augmented generation for multilingual long document understanding*. *arXiv:2508.07493*.
[11] Chen, J., Ma, W., Liu, P., Wang, W., Song, T., Li, M., **Wang, C.**, Qin, J., & Zhang, R., Chen, C. (2025). *MusiXQA: Advancing visual music understanding in multimodal large language models*. *arXiv:2506.23009*.
[12] Li, M., Zhang, R., **Wang, C.**, Chen, J., Gu, J., Zhou, Y., Dernoncourt, F., Zhu, W., Zhou, T., & Sun, T. (2025). *Towards visual text grounding of multimodal large language model*. *arXiv:2504.04974*.

Technical Reports & Datasets

[1] Do, T., Wang, W., Amini, M., Abdelhady, A., Xu, S., Negri, R., Kameshwar, S., **Wang, C.**, Dang, H., Jana, A., Carter, E., Alam, M., Kijewski-Correa, T., Prevatt, D., Roueche, D., & Wolohan, K. (2025). *StEER: Hurricane Milton Preliminary Virtual Reconnaissance Report (PVRR)*. doi:10.17603/ds2-resg-ah65.
[2] García Mejía, S., Jana, A., Erazo, K., **Wang, C.**, Diekmann, A., Xu, S., Romão, X., Kyprioti, A., Petreski, B., Bektaş, N., Lahna, T., Dang, H., Arora, P., Mostafa, K., & Wolohan, K. (2024). *2024 Hualien City Earthquake Media Repository*. doi:10.17603/ds2-4xv5-qc41.
[3] Alhawamdeh, B., Hassan, W., Gunay, S., Mosalam, K., Vargas, L., Marinković, M., Archbold, J., Martin, A., Merino-Peña, Y., Lahna, T., Hamdouni Alami, M., Burton, H., Iturburu, L., Ceferino, L., Duran, B., Nobahar, M., Romão, X., **Wang, C.**, Zhou, G., & Bektaş, N. (2023). *StEER: 2023 Mw 6.8 Oukaïmedene Morocco Preliminary Virtual Reconnaissance Report (PVRR)*. doi:10.17603/ds2-gw0j-6757.

Service & Outreach

Reviewer Service: ACL, EMNLP, NAACL, ICLR, NeurIPS, IJDRR
Lectures:

- *Data Visualization in Geographic Information Systems*, CIV 335: Data Analytics for Civil Engineering Systems, Stony Brook University (Spring 2024)
- *Applications of Data Analytical Methods in Highway Design*, CIV 305: Transportation System Analysis, Stony Brook University (Fall 2024)

Invited Talks:

- *How Advanced AI Reforms the Future of Civil Engineering*, Invited Seminar Talk, Stony Brook University (2025)

Awards & Honors

Provost Doctoral Fellowship, Stevens Institute of Technology (2022)
ECE Research Scholarship Award, Stevens Institute of Technology (2022)

Skills

Programming & Development: *Python, C/C++, SQL, Git, Linux, Docker*.
NLP & ML Frameworks: *PyTorch, TensorFlow, HuggingFace Transformers, scikit-learn, spaCy, FAISS*.
Data Processing & Visualization: *Pandas, NumPy, Matplotlib, Seaborn, GIS (ArcGIS/QGIS), Tableau*.
MLOps & Cloud: *AWS, GCP, MLflow, Weights & Biases, Airflow*.

Representative Repositories:

- *CrisisNLP*: Crowdsourced disaster response NLP pipeline; PyTorch, Transformers, In-Context Learning, Real-time data processing.
- *CaughtCheating*: MLLM benchmark for visual perception and reasoning; PyTorch, Multimodal LLM Inference & Evaluation.
- *Mosaic-IT*: Instruction tuning with compositional data synthesis; HuggingFace Transformers, LLM instruction-tuning, .
- *RuleR*: Rule-based data recycling for LLM controllability; PyTorch, Data pipeline design, spaCy.