

Yunsheng Ma

✉ yunsheng@purdue.edu | 🏠 ysma.me | 🎓 [Google Scholar](https://scholar.google.com/citations?user=...) | [in linkedin.com/in/yunsheng-ma](https://www.linkedin.com/in/yunsheng-ma)

EDUCATION

Purdue University PhD, Research Focus: Autonomous Driving and Foundation Models	West Lafayette, IN Jan. 2023 – Present
New York University Master of Science, Computer Science	New York City, NY Sep. 2020 – May 2022
Harbin Institute of Technology Bachelor of Engineering, Computer Software Engineering	Weihai, China Sep. 2016 – May 2020
University of California, Berkeley Undergraduate Exchange Student, Electrical Engineering and Computer Sciences	Berkeley, CA Aug. 2018 – May 2019

PUBLICATIONS

†denotes co-first authors

- [1] **Y. Ma**[†], C. Cui[†], X. Cao[†], W. Ye, P. Liu, J. Lu, A. Abdelraouf, R. Gupta, K. Han, A. Bera, J. M. Rehg, and Z. Wang. “LaMPilot: An Open Benchmark Dataset for Autonomous Driving with Language Model Programs.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [2] X. Cao[†], T. Zhou[†], **Y. Ma**[†], W. Ye, C. Cui, K. Tang, Z. Cao, K. Liang, Z. Wang, J. M. Rehg, and C. Zheng. “MAPLM: A Real-World Large-Scale Vision-Language Benchmark for Map and Traffic Scene Understanding.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [3] J. Lu, C. Cui, **Y. Ma**, A. Bera, and Z. Wang. “Quantifying Uncertainty in Motion Prediction with Variational Bayesian Mixture.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [4] **Y. Ma**, X. Cao, W. Ye, C. Cui, K. Mei, and Z. Wang. “Learning Autonomous Driving Tasks via Human Feedbacks with Large Language Models.” In *Findings of the Association for Computational Linguistics: EMNLP*, 2024.
- [5] C. Cui, Z. Yang, Y. Zhou, **Y. Ma**, J. Lu, and Z. Wang. “Large Language Models for Autonomous Driving: Real-World Experiments.” In *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2024.
- [6] **Y. Ma**, J. Lu, C. Cui, S. Zhao, X. Cao, W. Ye, and Z. Wang. “MACP: Efficient Model Adaptation for Cooperative Perception.” In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024.
- [7] C. Cui[†], **Y. Ma**[†], X. Cao[†], W. Ye[†], Y. Zhou, K. Liang, J. Chen, J. Lu, Z. Yang, K. Liao, T. Gao, E. Li, K. Tang, Z. Cao, T. Zhou, A. Liu, X. Yan, S. Mei, J. Cao, Z. Wang, and C. Zheng. “A Survey on Multimodal Large Language Models for Autonomous Driving.” In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, 2024.
- [8] C. Cui, **Y. Ma**, X. Cao, W. Ye, and Z. Wang. “Drive As You Speak: Enabling Human-Like Interaction With Large Language Models in Autonomous Vehicles.” In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, 2024.
- [9] C. Cui[†], **Y. Ma**[†], X. Cao, W. Ye, and Z. Wang. “Receive, Reason, and React: Drive as You Say, With Large Language Models in Autonomous Vehicles.” *IEEE Intelligent Transportation Systems Magazine*, 2024.
- [10] **Y. Ma**, R. Du, A. Abdelraouf, K. Han, R. Gupta, and Z. Wang. “Driver Digital Twin for Online Recognition of Distracted Driving Behaviors.” *IEEE Transactions on Intelligent Vehicles*, 2024.

- [11] C. Cui, **Y. Ma**, J. Lu, and Z. Wang. “REDFormer: Radar Enlightens the Darkness of Camera Perception with Transformers.” *IEEE Transactions on Intelligent Vehicles*, 2023.
- [12] W. Ye, **Y. Ma**, X. Cao, and K. Tang. “Mitigating Transformer Overconfidence via Lipschitz Regularization.” In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence (UAI)*, 2023.
- [13] **Y. Ma**, W. Ye, X. Cao, A. Abdelraouf, K. Han, R. Gupta, and Z. Wang. “CEMFormer: Learning to Predict Driver Intentions from In-Cabin and External Cameras via Spatial-Temporal Transformers.” *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2023.
- [14] C. Cui, **Y. Ma**, J. Lu, and Z. Wang. “Radar Enlighten the Dark: Enhancing Low-Visibility Perception for Automated Vehicles with Camera-Radar Fusion.” *IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.
- [15] **Y. Ma**, L. Yuan, A. Abdelraouf, K. Han, R. Gupta, Z. Li, and Z. Wang. “M²DAR: Multi-View Multi-Scale Driver Action Recognition With Vision Transformer.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2023.
- [16] L. Yuan, **Y. Ma**, L. Su, and Z. Wang. “Peer-to-Peer Federated Continual Learning for Naturalistic Driving Action Recognition.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2023.
- [17] **Y. Ma**, and Z. Wang. “ViT-DD: Multi-Task Vision Transformer for Semi-Supervised Driver Distraction Detection.” *IEEE Intelligent Vehicles Symposium*, 2022.
- [18] S. Zhao[†], **Y. Ma**[†], Y. Gu, J. Yang, T. Xing, P. Xu, R. Hu, H. Chai, and K. Keutzer. “An End-to-End Visual-Audio Attention Network for Emotion Recognition in User-Generated Videos.” In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2020, Oral Presentation.

PREPRINTS

- [1] **Y. Ma**, A. Abdelraouf, R. Gupta, Z. Wang, and K. Han. “Video Token Sparsification for Efficient Multimodal LLMs in Autonomous Driving.” *arXiv*, 2024.
- [2] W. Ye, G. Zheng, **Y. Ma**, X. Cao, B. Lai, J. M. Rehg, and A. Zhang. “MM-SpuBench: Towards Better Understanding of Spurious Biases in Multimodal LLMs.” *NeurIPS 2024 Workshop on Responsibly Building the Next Generation of Multimodal Foundational Models*, 2024.
- [3] X. Cao, B. Lai, W. Ye, **Y. Ma**, J. Heintz, J. Chen, J. Cao, and J. M. Rehg. “What is the Visual Cognition Gap between Humans and Multimodal LLMs?.” *arXiv*, 2024.
- [4] W. Ye, G. Zheng, X. Cao, **Y. Ma**, and A. Zhang. “Spurious Correlations in Machine Learning: A Survey.” *ICML 2024 Workshop on Data-Centric Machine Learning Research*, 2024.

WORK EXPERIENCE

Bosch Center for Artificial Intelligence

Research Intern, Supervisor: Dr. Burhaneddin Yaman

- Foundation Models for Autonomous Planning

Sunnyvale, CA

Sep. 2024 - Present

Toyota North America

Research Intern, Supervisor: Dr. Amr Abdelraouf

- Efficient Multimodal LLMs for Autonomous Driving

Mountain View, CA

May 2024 - Aug. 2024

Purdue University

Graduate Research Assistant, Supervisor: Prof. Ziran Wang

- Foundation Models for Autonomous Driving

West Lafayette, IN

Aug. 2022 – Present

- Proposed a LLM-based framework to enable human-centered decision-making in autonomous driving, featuring a language model program (LMP)-based planner that effectively learns from and adapt to human feedback (*accepted at CVPR 2024 and EMNLP 2024 Findings as first author*).

- Proposed a vision-language benchmark for autonomous driving, enhancing cross-modal traffic scene understanding (*accepted at CVPR 2024 as first author*).

- BEV-Based 3D Perception

- Proposed a parameter-efficient fine-tuning (PEFT) framework that adapts pre-trained single-agent models for cooperative perception (*accepted at WACV 2024 as first author*)

- Vision-Based Driver Monitoring

- Proposed a human digital twin framework that exploits transformer models for driver action recognition and temporal action localization. (*accepted at ITSC 2023 and IEEE T-IV as first author*)

New York University

Graduate Teaching Assistant

New York, NY

Spring 2021

- CSCI-GA.3033 Design and Analysis of Algorithms

DiDi

Computer Vision Research Intern, Supervisor: Dr. Pengfei Xu

Beijing, China

June 2019 – Sep. 2019

- Video Understanding

- Proposed a visual-audio attention network for end-to-end video emotion recognition that integrates spatial, channel-wise, and temporal attentions in 3D CNNs. (*accepted at AAI 2020 as first author*)

PROFESSIONAL ACTIVITIES

- As a Workshop Organizer

- WACV 2025 Workshop on Large Language and Vision Models for Autonomous Driving ([3rd LLVM-AD](#))
- ITSC 2024 Workshop on Large Language and Vision Models for Autonomous Driving ([2nd LLVM-AD](#))
- WACV 2024 Workshop on Large Language and Vision Models for Autonomous Driving ([1st LLVM-AD](#))

- As a Reviewer (77 papers in total)

CVPR 2024, ECCV 2024, WACV 2025, ICLR 2025, ICML 2024, AISTATS 2025, AAI 2025, IJCAI 2024, ITSC 2023/24, IV 2023/24, FISTS 2024, ICRA 2025, IROS 2024, ISBI 2024, CIKM 2024, ICCPS 2023, MOST 2023, IEEE IoT-J, IEEE T-IV, Taylor & Francis IJHCI, IEEE VTM, IEEE IC

- As a Volunteer

- Student Volunteer: UAI 2023, AAI 2023, TRB ITAP 2023
- Webmaster: IEEE Technical Committee on Internet of Things in Intelligent Transportation System

SELECTED AWARDS

IV Workshop on Foundation Intelligence for Vehicles: Best Paper Award	2024
NGTS Outstanding Speaker Award	2023
AAAI Travel Grant	2023
NeurIPS ML4AD Grant	2022
National Scholarship (top 1%)	2017

TECHNICAL SKILLS

Programming: Python, C++

Libraries: PyTorch, Transformers, Lightning, LangChain, MMDetection3D, OpenCV

Tools: CARLA, Chroma, Git, \LaTeX , SQL, Wandb