Fan Ma

Postdoctoral Associate, Department of Biomedical Informatics and Data Science, Yale University

flowerfan524@gmail.com flowerfan.site Google Scholar github.com/flowerfan

Research

I develop healthcare foundation models trained on longitudinal medical claims, electronic health records and physiological signals, and safety-oriented clinical reasoning systems built on large language models. My earlier research covered learning from limited and noisy supervision, temporal and video modeling, vision–language models and generative models.

Google Scholar: 2,278 citations, h-index 23, i10-index 31 (November 2026).

Positions

Postdoctoral Associate, Biomedical Informatics and Data Science, Yale University Nov 2024 – present
Advisor: Prof. Hua Xu. Healthcare foundation models on nationwide claims and EHR data; clinical reasoning with LLMs.

Postdoctoral Researcher, Zhejiang University Dec 2022 – Nov 2024
Multimodal learning, video understanding and generative models.

Research Intern, ByteDance Apr 2022 – Dec 2022
Video–language pre-training and video language models.

Research Intern, Meta (Facebook AI) May 2021 – Dec 2021
Unified transformer models for object tracking.

Research Intern, Baidu Research May 2019 – Dec 2019

Research Assistant, The Chinese University of Hong Kong Feb 2017 – Dec 2017

Education

Ph.D., Computer Science, University of Technology Sydney 2018 – 2022
Thesis: Learning from Imperfect Supervision in Visual Pattern Classification and Localization. Advisor: Prof. Yi Yang.

M.Eng., Xi'an Jiaotong University 2014 – 2017
Advisor: Prof. Deyu Meng.

B.Eng., Xi'an Jiaotong University 2010 – 2014

Publications

* equal contribution. Additional computer vision and language-model preprints are listed on Google Scholar.

Healthcare AI manuscripts

F. Ma*, M. Giuffrè*, D. Wright, K. McCann, M. Iscoe, L. Qian, M. Jiang, C. W. Ng, …, H. Xu. **A Safety-Oriented Hypothetico-Deductive Framework for AI-Assisted Differential Diagnosis.** arXiv preprint, 2026 (arXiv:2607.08038).

F. Ma, Q. An, P. Chen, L. Qian, X. Lan, M. Jiang, Z. Gu, X. Papademetris, H. Xu. **TRACE: Temporal Routing with Autoregressive Cross-channel Experts for EEG Representation Learning.** arXiv preprint, 2026 (arXiv:2605.11380).

F. Ma*, Y. Liu*, X. Lan*, W. Zhou, J. Ni, M. Giuffrè, L. Qian, X. Peng, …, H. Xu. **Foundation Models to Unlock Real-World Evidence from Nationwide Medical Claims.** arXiv preprint, 2026 (arXiv:2605.02740).

M. Jiang, Y. Li, D. Moyer, **F. Ma**, H. Xu, C. Chang. **Dataset-Dependent Effects of Cross-Depth Aggregation and Soft-Routed Experts in EEG Foundation Model Fine-Tuning.** arXiv preprint, 2026 (arXiv:2609.17886).

M.-H. Nguyen, E. Gray, B. A. Schuler, K. W. Byram, C.-T. Yang, **F. Ma**, H. Xu, W.-C. Su, …, C. Shyr. **Teaching Agentic AI to Generalize Expert Diagnostic Reasoning in Rare Diseases (LiteOdyssey).** arXiv preprint, 2026 (arXiv:2606.16149).

L. Qian, M. Giuffrè, Y. Wang, H. He, Q. Xie, X. Ai, X. Peng, **F. Ma**, …, H. Xu. **EHRNavigator: A Multi-Agent System for Patient-Level Clinical Question Answering over Heterogeneous Electronic Health Records.** arXiv preprint, 2026 (arXiv:2601.10020).

H. Kim, Y. Yin, Z. Cao, C. Liu, A. Li, Z. Chen, X. Ai, Y. Chung, **F. Ma**, X. Peng, et al. **How Far Have Large Language Models Advanced in Ophthalmology? A Systematic Review of Their Development, Evaluation, and Readiness for Clinical Use.** Research Square preprint, 2026.

Journal articles

M.-H. Nguyen, C.-T. Yang, T. A. Cassini, **F. Ma**, R. Hamid, L. A. Bastarache, J. F. Peterson, H. Xu, L. Li, S. Ma, C. Shyr. **Diagnostic Accuracy of Large Language Models for Rare Diseases: A Systematic Review and Meta-Analysis.** *npj Digital Medicine*, 2026.

Y. Yang, R. Zhang, **F. Ma**. **Empowering Personalized and Privacy-Preserving for Image Super-Resolution with Decentralized Collaboration.** *ACM Transactions on Multimedia Computing, Communications and Applications*, 2025.

D. Zhou, Y. Li, **F. Ma**, Z. Yang, Y. Yang. **MIGC++: Advanced Multi-Instance Generation Controller for Image Synthesis.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025 (arXiv:2407.02329).

X. He, S. Chen, **F. Ma**, Z. Huang, X. Jin, Z. Liu, D. Fu, Y. Yang, J. Liu, J. Feng. **VLAB: Enhancing Video Language Pre-training by Feature Adapting and Blending.** *IEEE Transactions on Multimedia*, 2025 (arXiv:2305.13167).

F. Ma, L. Zhu, Y. Yang. **Weakly Supervised Moment Localization with Decoupled Consistent Concept Prediction.** *International Journal of Computer Vision*, 2022.

F. Ma, Y. Wu, X. Yu, Y. Yang. **Learning with Noisy Labels via Self-Reweighting from Class Centroids.** *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

F. Ma, D. Meng, X. Dong, Y. Yang. **Self-paced Multi-view Co-training.** *Journal of Machine Learning Research*, 2020.

X. Dong, L. Zheng, **F. Ma**, Y. Yang, D. Meng. **Few-Example Object Detection with Model Communication.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019 (arXiv:1706.08249).

Conference papers

Z. Zhou, **F. Ma**, C. Gui, X. Xia, H. Fan, Y. Yang, T.-S. Chua. **AnchorFlow: Training-Free 3D Editing via Latent Anchor-Aligned Flows.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026, Oral (arXiv:2511.22357).

Y. Li, D. Zhou, **F. Ma**, F. Li, D. He, Y. Yang. **FoleyDirector: Fine-Grained Temporal Steering for Video-to-Audio Generation via Structured Scripts.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026, Highlight (arXiv:2603.19857).

C. Xia, **F. Ma**, R. Quan, Y. Xu, K. Zhan, Y. Yang. **Echoes of Ownership: Adversarial-Guided Dual Injection for Copyright Protection in MLLMs.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026 (arXiv:2602.18845).

R. Xu, D. Zhou, **F. Ma**, Y. Yang. **ContextGen: Contextual Layout Anchoring for Identity-Consistent Multi-Instance Generation.** *International Conference on Learning Representations (ICLR)*, 2026 (arXiv:2510.11000).

C. Meng, **F. Ma**, C. Zhang, J. Miao, Y. Yang, Y. Zhuang. **LogiStory: A Logic-Aware Framework for Multi-Image Story Visualization.** *International Conference on Learning Representations (ICLR)*, 2026 (arXiv:2603.28082).

Z. Zhou, X. Xia, **F. Ma**, H. Fan, Y. Yang, T.-S. Chua. **DreamDPO: Aligning Text-to-3D Generation with Human Preferences via Direct Preference Optimization.** *International Conference on Machine Learning (ICML)*, 2025 (arXiv:2502.04370).

W. Zhuo, **F. Ma**, H. Fan. **InfiniDreamer: Arbitrarily Long Human Motion Generation via Segment Score Distillation.** *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025 (arXiv:2411.18303).

Y. Suo, **F. Ma**, L. Zhu, T. Wang, F. Rao, Y. Yang. **From Trial to Triumph: Advancing Long Video Understanding via Visual Context Sample Scaling and Self-reward Alignment.** *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025 (arXiv:2503.20472).

Z. Zhou, **F. Ma**, H. Fan, T.-S. Chua. **Zero-1-to-A: Zero-Shot One Image to Animatable Head Avatars Using Video Diffusion.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025 (arXiv:2503.15851).

Y. Li, **F. Ma**, Y. Yang. **Imagine and Seek: Improving Composed Image Retrieval with an Imagined Proxy.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025 (arXiv:2411.16752).

Y. Suo, **F. Ma**, K. Shen, L. Zhu, Y. Yang. **Long-horizon Visual Instruction Generation with Logic and Attribute Self-reflection.** *International Conference on Learning Representations (ICLR)*, 2025 (arXiv:2503.13500).

Z. Tian, R. Quan, **F. Ma**, K. Zhan, Y. Yang. **BrainGuard: Privacy-Preserving Multisubject Image Reconstructions from Brain Activities.** *AAAI Conference on Artificial Intelligence (AAAI)*, 2025, Oral (arXiv:2501.14309).

C. Meng, **F. Ma**, J. Miao, C. Zhang, Y. Yang, Y. Zhuang. **Image Regeneration: Evaluating Text-to-Image Model via Generating Identical Image with Multimodal Large Language Models.** *AAAI Conference on Artificial Intelligence (AAAI)*, 2025 (arXiv:2411.09449).

H. Miao, **F. Ma**, R. Quan, K. Zhan, Y. Yang. **Autonomous LLM-Enhanced Adversarial Attack for Text-to-Motion.** *AAAI Conference on Artificial Intelligence (AAAI)*, 2025 (arXiv:2408.00352).

F. Ma, X. Jin, H. Wang, Y. Xian, J. Feng, Y. Yang. **Vista-LLaMA: Reducing Hallucination in Video Language Models via Equal Distance to Visual Tokens.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2312.08870).

F. Ma, X. Jin, H. Wang, J. Huang, L. Zhu, Y. Yang. **Stitching Segments and Sentences towards Generalization in Video-Text Pre-training.** *AAAI Conference on Artificial Intelligence (AAAI)*, 2024 (arXiv:2301.07463).

D. Zhou, Y. Li, **F. Ma**, X. Zhang, Y. Yang. **MIGC: Multi-Instance Generation Controller for Text-to-Image Synthesis.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2402.05408).

C. Liang, **F. Ma**, L. Zhu, Y. Deng, Y. Yang. **CapHuman: Capture Your Moments in Parallel Universes.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2402.00627).

R. Quan, W. Wang, Z. Tian, **F. Ma**, Y. Yang. **Psychometry: An Omnifit Model for Image Reconstruction from Human Brain Activity.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2403.20022).

R. Quan, W. Wang, **F. Ma**, H. Fan, Y. Yang. **Clustering for Protein Representation Learning.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2404.00254).

Y. Suo, **F. Ma**, L. Zhu, Y. Yang. **Knowledge-Enhanced Dual-stream Zero-shot Composed Image Retrieval.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

T. Feng, W. Wang, **F. Ma**, Y. Yang. **LSK3DNet: Towards Effective and Efficient 3D Perception with Large Sparse Kernels.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024 (arXiv:2403.15173).

Z. Zhou, **F. Ma**, H. Fan, Y. Yang. **HeadStudio: Text to Animatable Head Avatars with 3D Gaussian Splatting.** *European Conference on Computer Vision (ECCV)*, 2024 (arXiv:2402.06149).

W. Zhuo, **F. Ma**, H. Fan, Y. Yang. **VividDreamer: Invariant Score Distillation for Hyper-Realistic Text-to-3D Generation.** *European Conference on Computer Vision (ECCV)*, 2024 (arXiv:2407.09822).

Y. Feng, **F. Ma**, W. Lin, C. Yao, J. Chen, Y. Yang. **FedPAM: Federated Personalized Augmentation Model for Text-to-Image Retrieval.** *ACM International Conference on Multimedia Retrieval (ICMR)*, 2024.

F. Ma, M. Z. Shou, L. Zhu, H. Fan, Y. Xu, Y. Yang, Z. Yan. **Unified Transformer Tracker for Object Tracking.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022 (arXiv:2203.15175).

F. Ma, L. Zhu, Y. Yang, S. Zha, G. Kundu, M. Feiszli, Z. Shou. **SF-Net: Single-Frame Supervision for Temporal Action Localization.** *European Conference on Computer Vision (ECCV)*, 2020 (arXiv:2003.06845).

H. Wang, C. Deng, **F. Ma**, Y. Yang. **Context Modulated Dynamic Networks for Actor and Action Video Segmentation with Language Queries.** *AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

F. Ma, H. Yang, H. Yin, X. Huang, C. Yan, X. Meng. **Online Learning to Rank in a Listwise Approach for Information Retrieval.** *IEEE International Conference on Multimedia and Expo (ICME)*, 2019.

F. Ma, D. Meng, Q. Xie, Z. Li, X. Dong. **Self-Paced Co-training.** *International Conference on Machine Learning (ICML)*, 2017.

X. Dong, D. Meng, **F. Ma**, Y. Yang. **A Dual-Network Progressive Approach to Weakly Supervised Object Detection.** *ACM International Conference on Multimedia (ACM MM)*, 2017.

Y. Xu, **F. Ma**, D. Meng, C. Ren, Y. Leung. **A Co-training Approach to the Classification of Local Climate Zones with Multi-source Data.** *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017.

Workshop papers

C. Meng, **F. Ma**, C. Zhang, J. Miao, Y. Yang, Y. Zhuang. **Online Micro-Gesture Recognition in Long Videos via Spatiotemporal Feature Encoding and Query-Based Temporal Detection.** *MiGA Workshop at IJCAI*, 2025.

Professional service

100 peer reviews and review rounds (2022–2026).

Conference reviewer: CVPR (2023–2026), ICCV 2025, ECCV 2026, NeurIPS (2023, 2025), ICLR (2024–2026), ICML 2024, AAAI (2025–2027).

Journal reviewer: IEEE Transactions on Image Processing; IEEE Transactions on Multimedia; IEEE Transactions on Circuits and Systems for Video Technology; Journal of Biomedical Informatics; ACM Transactions on Multimedia Computing, Communications and Applications; Knowledge-Based Systems; Neural Networks; Neurocomputing; Frontiers of Computer Science; Image and Vision Computing.