Wufei Ma

wufeim@gmail.com · https://wufeim.github.io

EDUCATION

Sep 2022 –   **Johns Hopkins University** – Baltimore, MD
Ph.D. in Computer Scicence advised by Prof. Alan Yuille. GPA: 4.0/4.0.

Sep 2021 –   **Purdue University** – West Lafayette, IN
May 2022   Graduate study in Computer Science. GPA: 4.0/4.0.

Jan 2017 –   **Rensselaer Polytechnic Institute** – Troy, NY
May 2020   B.S. in Computer Science & B.S. in Mathematics
Summa Cum Laude. GPA: 3.96/4.0. Dean's Honor List in every semester.
Outstanding Performance Award $\times$ 2.

INTERNSHIP EXPERIENCES

Jun 2025 –   **Amazon FAR (Frontier AI for Robotics)** – San Francisco, CA
Applied Scientist Intern. Supervisor: Liang-Chieh (Jay) Chen.
Visual tokenizer for scene understanding and generation.

Nov 2024 –   **Google Research** – Remote
May 2025   Student researcher. Supervisors: Shun Liao, Yanqi Zhou, Zhuoyuan Chen.
Self-supervised video pretraining.

May 2023 –   **Meta Reality Labs** – Burlingame, CA
Nov 2023   Research scientist intern. Supervisors: Kai Li, Huiyu Wang.
Video-text pretraining; evaluation of video diffusion models.

May 2022 –   **Amazon AWS AI** – Santa Clara, CA
Aug 2022   Applied scientist intern. Supervisors: Srikar Appalaraju, R Manmatha.
Visual pretraining for scene-text VQA.

Jan 2021 –   **Microsoft Research Asia (MSRA)** – Beijing, China
Aug 2021   Research intern. Supervisors: Bin Li, Jiahao Li.
Deep learning-based video compression.

Aug 2020 –   **Megvii (Face++) Research** – Beijing, China
Dec, 2020   Research intern. Supervisor: Zhikang Liu.
Monocular 3D object detection with occlusion reasoning.

AWARDS AND SCHOLARSHIPS

2021   "Stars of Tomorrow" Award, *Microsoft Research Asia*

2020   Award of Excellence, *Megvii Research*

2019   Outstanding Performance Award, *Prof. Lirong Xia (RPI)*

2018   Outstanding Performance Award, *Prof. David Goldschmidt (RPI)*

2015   National Undergraduate Scholarship, *Wuhan University*

Publications

[25] <u>Wufei Ma</u>*, Yu-Cheng Chou*, Qihao Liu*, Xingrui Wang, Celso M de Melo, Jianwen Xie, and Alan Yuille. **SpatialReasoner: Towards Explicit and Generalizable 3D Spatial Reasoning**. In *Advances in Neural Information Processing Systems (NeurIPS)*. 2025.

[24] <u>Wufei Ma</u>, Haoyu Chen, Guofeng Zhang, Celso M de Melo, and Alan Yuille. **3DSRBench: A Comprehensive 3D Spatial Reasoning Benchmark**. In *International Conference on Computer Vision (ICCV)*. 2025.

[23] <u>Wufei Ma</u>, Luoxin Ye, Nessa McWeeney, Celso M de Melo, Jieneng Chen, and Alan Yuille. **A Compound 3D-Informed Design toward Spatially-Intelligent Large Multimodal Models**. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2025. **(Highlight presentation, 3%)**.

[22] Xingrui Wang, <u>Wufei Ma</u>, Tiezheng Zhang, Celso M de Melo, Jieneng Chen, and Alan Yuille. **Spatial457: A Diagnostic Benchmark for Comprehensive Spatial Reasoning of Large Multimodal Models**. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2025. **(Highlight presentation, 3%)**.

[21] Weijie Guo, Guofeng Zhang, <u>Wufei Ma</u>, and Alan Yuille. **DINeMo: Learning Neural Mesh Models with no 3D Annotations**. In *3rd Workshop on Compositional 3D Vision at CVPR*. 2025.

[20] Xingrui Wang, <u>Wufei Ma</u>, Angtian Wang, Shuo Chen, Adam Kortylewski, and Alan Yuille. **Compositional 4D Dynamics Scene Understanding with Physics Prior for Video Question Answering**. In *International Conference on Learning Representations (ICLR)*. 2025.

[19] <u>Wufei Ma</u>, Guofeng Zhang, Qihao Liu, Guanning Zeng, Adam Kortylewski, Yaoyao Liu, and Alan Yuille. **ImageNet3D: Towards General-Purpose Object-Level 3D Understanding**. In *Advances in Neural Information Processing Systems (NeurIPS)*. 2024.

[18] <u>Wufei Ma</u>, Kai Li, Zhongshi Jiang, Moustafa Meshry, Qihao Liu, Huiyu Wang, Christian Häne, and Alan Yuille. **Rethinking Video-Text Understanding: Retrieval from Counterfactually-Augmented Data**. In *European Conference on Computer Vision (ECCV)*. 2024. **(Strong double blind)**.

[17] Artur Jesslen, Guofeng Zhang, Angtian Wang, <u>Wufei Ma</u>, Alan Yuille, and Adam Kortylewski. **NOVUM: Neural Object Volumes for Robust Object Classification**. In *European Conference on Computer Vision (ECCV)*. 2024.

[16] <u>Wufei Ma</u>*, Qihao Liu*, Jiahao Wang*, Angtian Wang, Yaoyao Liu, Adam Kortylewski, and Alan Yuille. **Generating Images with 3D Annotations Using Diffusion Models**. In *International Conference on Learning Representations (ICLR)*. 2024. **(Spotlight presentation, 5%)**.

[15] <u>Wufei Ma</u>, Jiahao Li, Bin Li, and Yan Lu. **Uncertainty-Aware Deep Video Compression with Ensembles**. In *IEEE Transactions on Multimedia (TMM)*. 2024.

[14]   Bingchen Zhao, Jiahao Wang, <u>Wufei Ma</u>, Artur Jesslen, Siwei Yang, Shaozuo Yu, Oliver Zendel, Christian Theobalt, Alan Yuille, and Adam Kortylewski. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. 2024.

[13]   Angtian Wang*, <u>Wufei Ma</u>*, Alan Yuille, and Adam Kortylewski. **Neural Textured Deformable Meshes for Robust Analysis-and-Synthesis**. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 2023.

[12]   Jiahao Yang, <u>Wufei Ma</u>, Angtian Wang, Xiaoding Yuan, Adam Kortylewski, and Alan Yuille. **Robust Category-Level 3D Pose Estimation from Synthetic Data**. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 2023.

[11]   Xingrui Wang, <u>Wufei Ma</u>, Zhuowan Li, Adam Kortylewski, and Alan Yuille. **3D-Aware Visual Question Answering about Parts, Poses, and Occlusions**. In *Advances in Neural Information Processing Systems (NeurIPS)*. 2023.

[10]   Jiacong Xu, Yi Zhang, Jiawei Peng, <u>Wufei Ma</u>, ..., Alan Yuille, and Adam Kortylewski. **Animal3D: A Comprehensive Dataset of 3D Animal Pose and Shape**. In *IEEE/CVF International Conference on Computer Vision (ICCV)*. 2023.

[9]   Zhuowan Li, Xingrui Wang, Elias Stengel-Eskin, Adam Kortylewski, <u>Wufei Ma</u>, Benjamin Van Durme, and Alan Yuille. **SuperCLEVR: A Virtual Benchmark to Diagnose Robustness in Visual Reasoning**. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023. **(Highlight presentation)**

[8]   <u>Wufei Ma</u>, Angtian Wang, Adam Kortylewski, and Alan Yuille. **Robust Category-Level 6D Pose Estimation with Coarse-to-Fine Rendering of Neural Features**. In *European Conference on Computer Vision (ECCV)*. 2022.

[7]   Bingchen Zhao, Shaozuo Yu, <u>Wufei Ma</u>, Mingxin Yu, Shenxiao Mei, Angtian Wang, Ju He, Alan Yuille, and Adam Kortylewski. **OOD-CV: A Benchmark for Robustness to Out-of-Distribution Shifts of Individual Nuisances in Natural Images**. In *European Conference on Computer Vision (ECCV)*. 2022. **(Oral presentation)**

[6]   Xiaowei Zhang, <u>Wufei Ma</u>, Gunder Varinlioglu, Nick Rauh, Liu He, and Daniel Aliaga. **Guided Pluralistic Building Contour Completion**. 2022.

[5]   Farhad Mohsin, Lei Luo, <u>Wufei Ma</u>, Inwon Kang, Zhibing Zhao, Ao Liu, Rohit Vaish, and Lirong Xia. **Making group decisions from natural language-based preferences**. In *International Workshop on Computational Social Choice (COMSOC)*. 2021.

[4]   Arun Baskaran, Elizabeth J Kautz, Aritra Chowdhary, <u>Wufei Ma</u>, Bülen Yener, and Daniel Lewis. **Adoption of Image-Driven Machine Learning for Microstructure Characterization and Materials Design: A Perspective**. In *JOM, The Journal of The Minerals, Metals & Materials Society (TMS)*. 2021.

[3]   Elizabeth J Kautz, <u>Wufei Ma</u>, Arun Bakaran, Vineet Joshi, Bülent Yener, and Daniel Lewis. **Image-driven discriminative and generative methods for establishing microstructure-processing relationships relevant to nuclear fuel processing pipelines**. In *Microscopy and Microanalysis*. 2021.

[2]   <u>Wufei Ma</u>, Elizabeth J Kautz, Arun Bakaran, Aritra Chowdhary, Vineet Joshi, Bülent Yener, and Daniel Lewis. **Image-driven discriminative and generative machine learning algorithms for establishing microstructure–processing relationships**. In *Journal of Applied Physics*. 2020.

[1]   Elizabeth J Kautz, <u>Wufei Ma</u>, Saumyadeep Jana, Arun Devaraj, Vineet Joshi, Bülent Yener, and Daniel Lewis. **An image-driven machine learning approach to kinetic modeling of a discontinuous precipitation reaction**. In *Materials Characterization*, 2020.

PROFESSIONAL SERVICES

2021 –   **Reviewer**
ICLR, NeurIPS, ICML, CVPR, ICCV, ECCV, AAAI, WACV, etc.

TEACHING EXPERIENCES

Fall 2025   **Graduate teaching assistant**, *Johns Hopkins University*
AS675 - Probabilistic Models of the Visual Cortex.

Spring 2024   **Graduate course assistant**, *Johns Hopkins University*
CS661 - Computer Vision.

Fall 2023   **Graduate course assistant**, *Johns Hopkins University*
CS661 - Computer Vision.

Spring 2023   **Graduate course assistant**, *Johns Hopkins University*
CS671 - NLP: Self-Supervised Models.

LEADERSHIP

Sep 2015 –   **Captain**, *Soccer team of College of Mathematics at Wuhan University*
Nov 2016   Enter semi-final and quarter-final of WHU Soccer Champion Cup in 2015 and 2016, out of 32.

Jan 2014 –   **President**, *Rubik's Cube Club at Shanghai High School*
Jan 2015   Organize tutorials, workshops, and competitions for various Rubik's Cubes.