Guangyu Meng

gmeng@nd.edu | gm3g11.github.io | linkedin.com/in/guangyumeng | github.com/gm3g11 | (574) 340-6577

RESEARCH INTERESTS

My research develops efficient and interpretable methods at the intersection of machine learning, computational geometry, and topological data analysis. My current focus is on medical image analysis by integrating deep learning with persistent homology for self-supervised learning and clustering. I also work on computational topology, including Reeb graph comparison and optimal transport, as well as curriculum learning strategies for efficient large language model fine-tuning.

EDUCATION

University of Notre Dame — **South Bend, IN**
Ph.D. in Computer Science (GPA: 3.92/4.0) — *2021 – Present*
- Co-advisors: Prof. Erin W. Chambers, Prof. Danny Z. Chen

Washington University in St. Louis — **St. Louis, MO**
M.S. in Computer Science (GPA: 4.0/4.0) — *2020 – 2021*

University of Southampton — **Southampton, UK**
M.S. in Microelectronics System Design (With Merit) — *2011 – 2012*

Huazhong University of Science and Technology — **Wuhan, China**
B.E. in Electronic Science and Engineering (GPA: 3.7/4.0) — *2007 – 2011*

RESEARCH EXPERIENCE

Graduate Research Assistant — **University of Notre Dame**
Department of Computer Science and Engineering — *2021 – Present*

- **VTRL-Med – Medical Image Clustering:** Achieved state-of-the-art performance with 81.51% average accuracy across 13 medical benchmarks, representing a 4.48% improvement over existing baselines while nearly matching supervised methods in a fully unsupervised setting. Developed a novel visual-structural and topology-aware framework that adapts the Segment Anything Model (SAM) with mask-based structural priors and persistent homology features to enable anatomically consistent clustering of medical images.
- **TopoCL – Topology-Enhanced Contrastive Learning:** Demonstrated consistent accuracy improvements averaging +3.26% with strong statistical significance (80% of comparisons with $p < 0.001$) across five state-of-the-art contrastive learning methods (SimCLR, MoCo-v3, BYOL, DINO, and Barlow Twins) on five medical imaging datasets. Developed a universal framework featuring topology-aware augmentations that quantify and control topological perturbations, a Hierarchical Topology Encoder for capturing topological structures, and an adaptive mixture-of-experts fusion module for integrating visual and topological representations.
- **Reeb Graph Comparison – Computational Geometry:** Designed a stable algorithmic framework for comparing Reeb graphs using Gromov-Wasserstein distance with persistence images, accompanied by rigorous theoretical stability proofs under scalar field perturbations. This work provides theoretical guarantees for topological shape analysis and comparison in computational geometry applications.
- **PUDF – Curriculum Learning for Large Language Models:** Developed a psychology-based curriculum learning framework that applies Item Response Theory to artificial crowds for model-independent difficulty quantification and Dynamic Data Selection via Model Ability Estimation (DDS-MAE) for adaptive training scheduling. Demonstrated effectiveness on both classification and generation tasks: fine-tuning Llama3.1-8B on AG News (1 million examples) achieved 4.13% accuracy improvement and 69.68% faster training compared to baseline, with 0.42% accuracy gain and 75.48% speedup over state-of-the-art reinforcement learning curriculum methods; on GSM8K mathematical reasoning tasks, achieved 76.70% accuracy with Qwen2.5-7B (42% faster) and 61.72% with Llama3.1-8B (27% faster).
- **NNS-EMD – Efficient Optimal Transport:** Achieved 44× to 135× speedup over exact Earth Mover's Distance computation while maintaining superior accuracy on image and document classification and retrieval tasks (91.88% on 20news, 93.24% on Amazon reviews compared to 87.42% and 78.53% for ACT baseline). Proposed a GPU-accelerated Nearest Neighbor Search approximation algorithm with comprehensive theoretical analysis of time complexity and error bounds, enabling scalable optimal transport for large-scale computer vision applications.

PUBLICATIONS

- **Guangyu Meng**, Pengfei Gu, John P. Lalor, Peixian Liang, Erin W. Chambers, Danny Z. Chen, "VTRL-Med: Vision-Topology Reinforced Learning for Medical Image Clustering," under review at CVPR 2026.
- **Guangyu Meng**, Pengfei Gu, John P. Lalor, Peixian Liang, Erin W. Chambers, Danny Z. Chen, "TopoCL: Topology-Enhanced Contrastive Learning for Medical Image Analysis," under review at CVPR 2026.

- **Guangyu Meng**, Erin W. Chambers, "A Stable and Theoretically Grounded Gromov-Wasserstein Distance for Reeb Graph Comparison using Persistence Images," under review at Symposium on Computational Geometry (SOCG) 2026.
- **Guangyu Meng**, Qingkai Zeng, John P. Lalor, Hong Yu, "Psychology-based Unified Dynamic Framework for Curriculum Learning," Computational Linguistics Journal, 2025.
- Peixian Liang, Yifan Ding, Yizhe Zhang, Jianxu Chen, Hao Zheng, Hongxiao Wang, Yejia Zhang, **Guangyu Meng**, Tim Weninger, Michael Niemier, X. Sharon Hu, Danny Z. Chen, "Cell Instance Segmentation: The Devil Is in the Boundaries," IEEE Transactions on Medical Imaging (TMI), 2025.
- **Guangyu Meng**, Ruyu Zhou, Liu Liu, Peixian Liang, Fang Liu, Danny Z. Chen, Michael Niemier, X. Sharon Hu, "Efficient Approximation of Earth Mover's Distance Based on Nearest Neighbor Search," IEEE Transactions on Multimedia (TMM), 2025.
- **Guangyu Meng**, Qisheng Jiang, Kaiqun Fu, Beiyu Lin, Chang-Tien Lu, Zhiqian Chen, "Early Forecasting of the Impact of Traffic Accidents Using a Single Shot Observation," SIAM International Conference on Data Mining (SDM) 2022.

TEACHING EXPERIENCE

Teaching Assistant — **University of Notre Dame**
Department of Computer Science and Engineering — *2021 – Present*

- CSE 60111: Complexity and Algorithms, Spring 2025
- CSE 20110: Discrete Mathematics, Fall 2024
- CSE 40625/60625: Machine Learning, Fall 2021
- CSE 30321: Computer Architecture, Spring 2021

INDUSTRY EXPERIENCE

Emulation Specialist — **Mentor Graphics Co., Ltd**
Shenzhen, China — *2016 – 2019*

- Led high-performance verification solutions for Huawei's first commercial 5G chipset; optimized testing code, reducing runtime from 45 to 8 minutes through parallelization.
- Delivered technical presentations at Silicon Valley headquarters and led training for FiberHome on UVM environment configuration.

ASIC Design Engineer — **Spreadtrum Communications**
Shanghai, China — *2015 – 2016*

- Maintained chip-level emulation environments; developed Python-based resource allocation program improving device utilization by 20%.

NPI Engineer — **NXP Semiconductors**
Shanghai, China — *2013 – 2015*

- Optimized wafer mapping and packaging processes, enhancing die bond machine capacity by 14.8%.

TECHNICAL SKILLS

Programming Languages: Python, C/C++, CUDA, Java, MATLAB

Deep Learning: PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn, PEFT (LoRA, QLoRA)

Computer Vision: OpenCV, SAM (Segment Anything), DINO, Vision Transformers, ResNet, U-Net

Topological Data Analysis: Persistent Homology, Gudhi, Ripser, giotto-tda

Scientific Computing: NumPy, Pandas, SciPy, Matplotlib

Systems & Tools: Git, Docker, Linux/HPC (Slurm), CUDA Programming, GPU Optimization, Mixed-Precision Training

PROFESSIONAL SERVICE

Conference Reviewer: Symposium on Computational Geometry (SOCG) 2025, AAAI Conference on Artificial Intelligence 2024, SIAM International Conference on Data Mining (SDM) 2023