Jingwei Zhang

+1 (631) 310-7076
Jingwezhang@cs.stonybrook.edu
Google Scholar

EDUCATION

Stony Brook University Stony Brook, NY, USA
Ph.D. Candidate in Computer Science *2018.09 - now*
- Research interests: computer vision, efficient models, Mamba, GPU algorithms, multi-modal models, self supervised learning (SSL), parameter efficient fine tuning (PEFT)
- Advisor: Prof. Dimitris Samaras
- GPA: 4.00/4.00

Institute of Computing Technology, Chinese Academy of Sciences Beijing, China
M.S. in Computer Science *2015.09 - 2018.06*
- Research interests: algorithms in bioinformatics, statistics
- Advisor: Prof. Shiwei Sun and Prof. Dongbo Bu
- Major GPA: 3.89 / 4.00, Overall GPA: 3.73 / 4.00

Shandong University Shandong, China
B.E. in Computer Science and Technology *2010.09 - 2015.06*
- Major GPA: 3.94 / 4.00, Overall GPA: 3.81 / 4.00

EXPERIENCE

Insitro South San Francisco, CA, USA
Data Science and Machine Learning Intern *2025.06 - 2025.08*
- Explored CLIP alignment on cell images and their treatment text, proposed a many-to-one CLIP method to align multiple image and their text treatments with mixup to boost performance.
- Proposed a channel agnostic transformer: ChannelSFormer, which utilized a separated spatial/channel attention for multi-channel images and outperformed the current SOTA by 7% and is 3 times faster.

MAJOR RESEARCH PROJECTS

ChannelSFormer: A Channel-Agnostic Vision Transformer 2025.06 – 2026.01
- Engineered a novel channel-agnostic attention mechanism by decoupling global attention into independent channel-wise and spatial-wise components, optimizing both speed and model accuracy.
- Developed a channel class token mechanism that eliminates the requirement for fixed channel embeddings, enabling seamless adaptation to varying input channels and enhancing deployment flexibility in real-world applications.

LBMamba: Locally Bi-directional Mamba 2024.12-2025.05
- Proposed a Mamba architecture that employs a locally bi-directional scanning mechanism and its hardware-aware (thread-level bi-direction) selective scan CUDA operator. It has 27% more FLOPS than the vanilla Mamba but only 2% slower.
- Together with a sequence reserving operation, our model only need one scan operations every layer, achieving much better accuracy-throughput trade-off than the baseline methods.

2DMamba: Efficient State Space Model for Image Representation 2024.06-2024.11
- Proposed a 2D selective State Space Model architecture which directly scans a 2D image without first flattening it into a 1D sequence, which results in a loss of spatial correlation in 2D.
- Proposed a novel hardware-aware 2D selective scan CUDA operator with a special designed caching mechanism that extends the 1D Mamba parallelism into 2D.
- Demonstrated the effectiveness of our method over SOTA methods on both pathology classification/survival tasks and natural image classification/segmentation tasks.

Cell-Graph-Aware Self Supervised Learning for Histopathology Images 2023.09-2024.06

- Proposed a biological entity-aware SSL framework that aligns pixel-level features with topological structures via Graph Neural Networks (GNN) to capture complex tissue micro-environments.
- Validated on TCGA and BRIGHT datasets, demonstrating that GNN-enhanced image-graph features significantly outperform single-modality baselines in downstream tasks.

A Segment Anything Model for Semantic Segmentation in Digital Pathology 2023.05-2023.08
- Proposed SAM-Path, which adapted the Segment Anything Model (SAM) for semantic segmentation in pathology by introducing trainable class prompts, enabling SAM for semantic segmentation tasks.
- Further improved SAM's applicability to pathology tasks by incorporating a pathology encoder, specifically a pathology foundation model.
- Demonstrated the effectiveness of SAM-Path in conducting semantic segmentation in digital pathology without the need for human input prompts through extensive experiments on public pathology datasets.

Boosting Multi-Instance Learning Schemes via Task-specific Prompt Tuning 2023.01-2023.03
- Proposed Prompt-MIL, a novel multi-instance learning (MIL) framework for whole slide image (WSI) classification in computational pathology.
- Introduced a prompt tuning mechanism for MIL that fine-tunes a small fraction of parameters to inject task-specific knowledge into self supervised features.
- Conducted extensive experiments on three WSI datasets, demonstrating Prompt-MIL's superior accuracy over conventional MIL methods. Compared to traditional full fine-tuning approaches, our method is more accurate, fast and GPU memory saving.

Precise Location Matching Improves Dense Contrastive Learning 2022.09-2022.12
- Proposed a precise location-based matching mechanism to boost contrastive learning for dense prediction tasks in digital pathology.
- Leveraged overlapping information between geometric transformations to precisely match regions in two augmentations.
- Experiments on two pretraining datasets and three downstream datasets showed superiority of proposed method in semantic and instance segmentation tasks.

Gigapixel Whole-Slide Images Classification Using Locally Supervised Learning 2021.09-2022.03
- Proposed a locally supervised learning framework for classifying whole slide images (WSIs) that optimizes a network by dividing it into modules and leveraging some auxiliary models for local optimization.
- Introduced a novel Random Feature Reconstruction unit (RFR) to enhance the preservation of distinguishing features during training, improving the performance by 1% to 3%.
- Achieved significant computational efficiency by being 7 to 10 times faster than conventional methods, and the method required only 20% of the total GPU memory compared to end-to-end training approaches.

A Joint Spatial and Magnification-Based Attention Framework 2020.09 – 2021.04
- Developed a novel attention sampling strategy that jointly optimizes spatial localization and magnification selection, enabling efficient processing of Gigapixel Whole Slide Images (WSIs).
- Implemented an end-to-end framework that leverages down-sampled global context to generate attention maps for spatial-magnification probability distributions, facilitating targeted high-resolution patch extraction.
- Validated performance on two histopathology datasets, achieving state-of-the-art (SOTA) accuracy while significantly reducing computational overhead compared to exhaustive patch-based inference and automated magnification selection.

MAJOR PUBLICATIONS

[1] **Zhang Jingwei**, et al. "ChannelSFormer: A Channel Agnostic Vision Transformer for Multi-Channel Cell Painting Images," (Acceptted to Imagemics@NeurIPS 2025, ready to submit to ICML)

[2] **Zhang Jingwei**, et al. "LBMamba: Locally Bi-directional Mamba,", *Transactions on Machine Learning Research (TMLR)*, 2025

[3] Han Xi, **Zhang Jingwei**, et al. "GeoMaNO: Geometric Mamba Neural Operator for Partial Differential Equations," (co-first author, ready to resubmit)

[4] **Zhang Jingwei**, et al. "2DMamba: Efficient State Space Model for Image Representation with Applications on Giga-Pixel Whole Slide Image Classification," *CVPR*, 2025, 3583-3592.

[5] Shi zhan, **Zhang Jingwei**, et al. "Integrative Graph-Transformer Framework for Histopathology Whole Slide Image Representation and Classification, " *MICCAI*, 2024, 341-350

[6] Kapse Saarthak, Pati Pushpak, Das Srijan, **Zhang Jingwei**, et al. "SI-MIL: Taming Deep MIL for Self-Interpretability in Gigapixel Histopathology," *CVPR*, 2024.

[7] Miao Qiaomu, Graikos Alexandros, **Zhang Jingwei**, et al. "Diffusion-Refined VQA Annotations for Semi-Supervised Gaze Following," *ECCV*, 2024.

[8] Kapse Saarthak, Das Srijan, **Zhang Jingwei**, et al. "Attention De-sparsification Matters: Inducing diversity in digital pathology representation learning," *Medical Image Analysis*, 2024

[9] **Zhang Jingwei**, et al. "Prompt-MIL: Boosting Multi-instance Learning Schemes via Task-Specific Prompt Tuning," *MICCAI*, 2023, 624-634. (**Oral**)

[10] **Zhang Jingwei**, et al. "SAM-Path: A Segment Anything Model for Semantic Segmentation in Digital Pathology," *MICCAI Workshops*, 2023, 161-170. (**Oral**)

[11] **Zhang Jingwei**, et al. "Precise Location Matching Improves Dense Contrastive Learning in Digital Pathology," *IPMI*, 2023, 783-794.

[12] **Zhang Jingwei**, et al. "Gigapixel Whole-Slide Images Classification Using Locally Supervised Learning," *MICCAI*, 2022, 192-201. (**Oral**)

[13] **Zhang Jingwei**, et al. "A Joint Spatial and Magnification Based Attention Framework for Large Scale Histopathology Classification," *CVPR Workshops*, 2021, 3776-3784. (**Oral**, Best paper award)

AWARDS & INTELLECTUAL PROPERTY

- U.S. Patent No. 63/912,208 (related to ChannelSFormer, in application) Nov, 2025
- Best Paper Award, in CVMI 2021 at CVPR 2021 June, 2021
- Best Poster Award, in CCF Bioinformatics Conference 2016 Nov, 2016
- Bronze Medal of the 2011 ACM-ICPC Asia Regional Contest Nov, 2011

SKILLS

Python, Pytorch, Java, C/C++, MATLAB, MySQL