Yunfei Xie

• Email: xieyunfei01@gmail.com • Tel: +1-831-295-9285

EDUCATION

Huazhong University of Science and Technology   09/2021-06/2025
- Candidate for **Bachelor of Engineering** in **Artificial Intelligence**   GPA: 87.0/100
- **Scholarship & Honors**: Science and Technology Scholarship (09/2022) & Second Prize in the C Programming Competition (06/2022)

PUBLICATIONS

- **Yunfei Xie**, Cihang Xie, Alan Yuille, et al. **"From Pixels to Objects: A Hierarchical Approach for Part and Object Segmentation Using Local and Global Aggregation."** Accepted by *European Conference on Computer Vision (ECCV 2024)*
- **Yunfei Xie**, Ce Zhou, Jieru Mei, et al. **"Brain Tumor Segmentation Through Supervoxel Transformer."** Accepted by *IEEE International Symposium on Biomedical Imaging 2024 (IEEE ISBI 2024)*
- **Yunfei Xie,** Ce Zhou, Lang Gao, et al. **"MedTrinity-25M: A Large-scale Multimodal Dataset with Multigranular Annotations for Medicine."** Submitted to *the 38th Annual Conference on Neural Information Processing Systems (NeruIPS 2024),* (under review)
- **Yunfei Xie**, Yuyin Zhou, Alan Yuille, et al. **"Few-Shot Medical Image Segmentation via Supervoxel Transformer."** Submitted to *the 39th Annual AAAI Conference on Artificial Intelligence (AAAI 2024)*, (under review)
- Yuzhou Gong, **Yunfei Xie**, Hanlin Li, et al. **"AM-SCNet: Activation Mask Sparse Convolutional Network for Skeleton-based Action Recognition."** Accepted by *the 6th International Conference on Information Engineering & Information Science (ICIEIS 2024)*

RESEARCH EXPERIENCE

Research Intern, VLAA at University of California, Santa Cruz   12/2023-Present
MedTrinity-25M: A Large-scale Multimodal Dataset with Multigranular Annotations for Medicine

Advisor: **Prof. Yuyin Zhou**

- Introduced a comprehensive, large-scale multimodal dataset for medicine, covering over 25 million images across 10 modalities, with multigranular annotations for more than 65 diseases, including detailed ROIs
- Developed an automated pipeline for scaling multimodal data without paired text descriptions, utilizing domain-specific models to identify ROIs and prompting multimodal LLMs for enriched textual annotations
- Pretrained medical-specific MLLMs on this dataset with enriched annotations for various multimodal tasks, achieving state-of-the-art performance in tasks like medical VQA.
- **The dataset is open-source and has received over 100 stars on GitHub**
- One paper in submission to **NeurIPS 2024** as the co-first author and project leader

Few-shot Medical Image Segmentation via Supervoxel Transformer   Advisor: **Prof. Yuyin Zhou**

- Introduced SVFormer, the first 3D Transformer-based few-shot framework for medical image segmentation, utilizing supervoxel representations to preserve 3D semantic details while reducing feature redundancy
- Developed a novel Supervoxel Cross Attention (SCA) clustering method to extract flexible supervoxel representations, effectively addressing the high dimensionality challenges of 3D data in transformer

architectures
- Proposed a supervoxel-based prototypical segmentation technique that generates interpretable 3D prototypes by aligning supervoxels with target organs, leveraging the full volumetric context of medical images
- Demonstrated superior performance over state-of-the-art methods across three public datasets
- One paper in submission to **AAAI 2024** as the individual first author

Brain Tumor Segmentation Through Supervoxel Transformer　　　　Advisor: **Prof. Yuyin Zhou**
- Investigated two CNN-Transformer hybrid models within the BraTS-ISBI 2024 challenge to develop segmentation models with broad applicability across brain tumors
- Introduced a supervoxel Transformer to efficiently cluster similar voxels to improve the interpretability, incorporating a supervoxel cross-attention mechanism to iteratively refine assignments and features
- Trained the models on an augmented dataset and combined them via ensembling, proving their superiority of the Supervoxel Transformer in enhancing tumor area
- One paper accepted by **IEEE ISBI 2024** as the co-first author and project leader

Research Intern, CCVL Lab at Johns Hopkins University　　　　　　06/2023-11/2023
From Pixels to Objects: A Hierarchical Approach for Part and Object Segmentation Using Local and Global Aggregation　　　　　　　　　　　　　　　Advisor: **Prof. Alan Yuille**
- Presented a hierarchical transformer-based model for image segmentation that effectively links the segmentation of detailed parts to the broader context of object segmentation
- Introduced an efficient hierarchical representation that mimics the human visual system, significantly reducing the complexity inherent in Vision Transformers by addressing both local and global redundancies
- Developed a novel upsampling method based on hierarchical representation that effectively mitigates the blurring issues common to existing upsampling techniques
- Proposed a hierarchical superpixel representation and unified framework to solve multi-level segmentation
- Outperformed the state-of-the-art models by 2.8% and 1.5% on two widely used benchmarks
- One paper accepted by **ECCV 2024** as the individual first author

Research Assistant, Shoudong Han Research Group at HUST　　　　　12/2022-02/2023
AM-SCNet: Activation Mask Sparse Convolutional Network for Skeleton-based Action Recognition
　　　　　　　　　　　　　　　　　　　　　　　　　　　　Advisor: **Prof. Shoudong Han**
- Proposed the AM-SCNet, an efficient method to recognize actions by skeleton maps, to resolve the high computational costs in skeleton-based recognition tasks
- Introduced Activation Mask (AM) into sparse convolutions for dynamically filtering spatiotemporal action semantics with significant sparse features
- Conducted experiments to prove AM-SCNet's efficiency, where the method achieves 93.8%, 96.7% accuracy on the X-Sub and X-View benchmark of NTU 60, while reducing GFLOPs by 85.9%
- One paper accepted by **ICIEIS 2024**

TECHNICAL SKILLS
- Programming Language: Python, C, and C++
- Tools: Linux, PyTorch, and OpenCV