Graduate Research Plan

Research Title: Efficient Hybrid Architectures: Scaling Convolutional Nearest Neighbors to Medical Imaging Inverse Problems.

Summary of Graduate Research Plan:

Medical imaging inverse problems such as accelerated MRI reconstruction, limited-angle computed tomography, and sparse-view imaging require neural architectures that simultaneously capture local structural patterns and long-range semantic dependencies while maintaining computational efficiency. Recent work of Attention U-Net [1] demonstrates that self-attention mechanisms from transformers improve image segmentation by capturing long-range anatomical dependencies, while standard U-Net [2] convolutions efficiently capture local spatial structure. However, current approaches face a fundamental tradeoff between high quality reconstruction and computational cost. High quality reconstructions require either computationally expensive attention mechanisms or limited convolution-based methods. This creates a critical barrier to deploying advanced reconstruction algorithms in clinical settings where computational resources are constrained and inference speed is clinically relevant.

My senior honors thesis introduced Convolutional Nearest Neighbors (ConvNN), a novel framework that selects k-nearest features in learned feature space rather than aggregating over spatial neighborhoods. This work demonstrated that ConvNN can match the performance of attention-based architectures on image classification tasks while reducing computational overhead by 25 percent. However, this research remains limited to proof-of-concept demonstrations and does not explore the scalability of this framework to larger clinical applications. The core opportunity is to demonstrate that ConvNN scales to clinical medical imaging inverse problems, becoming the first practical validation of feature-space aggregation for medical reconstruction.

I propose to develop and validate hybrid branching architectures that integrate ConvNN with standard convolution. The architecture consists of two parallel pathways per layer: a local pathway using standard convolutional blocks to capture immediate spatial structure, and a global pathway implementing ConvNN modules that use norm-based k-nearest neighbor selection in learned feature space to capture long-range semantic relationships. This design avoids the quadratic complexity of attention ($O(n^2)$) by replacing softmax aggregation with efficient k-nearest neighbor selection ($O(n \log n)$), achieving 25 to 30 percent reduction in FLOPs while maintaining comparable reconstruction quality.

The research unfolds in three phases. First, I will validate the architecture on the public fastMRI [3] dataset containing knee MRI images with 4x and 8x acceleration. I will compare three architectures: baseline U-Net with standard convolution only, Attention U-Net combining attention and convolution, and the proposed hybrid branching ConvNN U-Net. This public benchmark phase establishes architectural advantages in a controlled setting before moving to clinical data.

Building on these results, I will validate the approach on real clinical MRI data. I will need to work with a medical school or hospital to access clinical MRI data. These MRI data will be used to train and validate the reconstruction network. Validation includes radiologist assessment on a five point diagnostic quality scale, robustness testing across different anatomies, and deployment feasibility studies on standard clinical hardware. If clinical results validate the approach, I will extend the architecture to additional medical inverse problems such as limited-angle CT reconstruction and sparse-view tomography. This will establish that feature-space geometry principles generalize beyond MRI and serve as fundamental design principles for efficient medical image reconstruction.

Intellectual Merit

This research advances knowledge in three distinct ways. First, by demonstrating that k-nearest neighbor aggregation in learned feature space outperforms both spatial convolution alone and softmax-based attention on clinically relevant tasks, this work provides evidence that feature-space geometry rather than spatial locality is fundamental to efficient multi-scale reasoning. This insight challenges conventional architectural design choices and has broad implications for computer vision and neural architecture design.

Second, this research demonstrates that computational efficiency and reconstruction quality need not be competing objectives. Current understanding treats these as inherent tradeoffs and successful validation would demonstrate that principled architectural design choices can achieve both simultaneously, advancing methodology for neural architecture design.

Third, this research provides the first demonstration that ConvNN is effective on clinical-scale inverse problems. This represents meaningful validation bridging theoretical architectural innovation with practical clinical deployment, establishing whether feature-space geometry is a fundamental principle underlying efficient medical image reconstruction.

Broader Impacts

By reducing computational requirements by 25 to 30 percent, hybrid ConvNN architectures enable deployment on standard hospital infrastructure, lowering barriers for community hospitals and resource-limited settings to adopt AI imaging. This particularly benefits facilities without access to state-of-the-art computing resources, advancing equity in medical AI deployment.

Faster and more efficient reconstruction directly benefits patients through shorter MRI scan times. A 4x acceleration with diagnostic-quality reconstruction could reduce typical scan times from 30 minutes to 7.5 minutes. This improvement is especially important for elderly patients, pediatric populations, and those with anxiety or motion disorders, for whom long immobilization is particularly challenging.

By designing computationally efficient reconstruction networks, this research enables advanced imaging tools in resource-constrained clinical settings, advancing equity in access to diagnostic imaging. Data scarcity and computational barriers disproportionately affect rare diseases and populations in low-resource regions; efficient architectures directly address both barriers simultaneously.

References:

[1] Oktay, Ozan, et al. “Attention u-net: Learning where to look for the pancreas.” arXiv preprint arXiv:1804.03999 (2018)

[2] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation.” International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015.

[3] Zbontar, Jure, et al. “fastMRI: An open dataset and benchmarks for accelerated MRI.” arXiv preprint arXiv:1811.08839 (2018)