
Enabling Medical Image and Signal Analysis for Resource-constrained Environments

Research Statement

Mohammed Yusuf Ansari
ma1@tamu.edu

My research vision is to "*develop efficient and effective computational methods to enable the analysis, diagnosis, and generation of medical images and signals, specifically for resource-constrained environment*" (e.g., hospitals and clinics in developing countries, war-torn regions, etc.). Growing up in India, I witnessed many of my relatives and immediate family members suffer from cardiovascular diseases (CVDs) and cancer, often misdiagnosed due to a lack of automated and objective analysis. The subjectivity led to delayed diagnoses and ineffective treatments, which tragically resulted in preventable deaths. These personal experiences have driven my commitment to developing systems that provide accurate, timely, and reliable diagnostics for cancer and cardiovascular conditions, regardless of the infrastructure available.

Given the current state of technology, I have worked on the foundation of my vision by creating lightweight and efficient neural networks for analyzing biomedical images and signals. While deep neural networks have surpassed expert-level performance in diagnosing various diseases, their integration into healthcare systems remains limited due to high computational demands. My current research addresses this problem by proposing lightweight deep learning (DL) frameworks with a focus on efficiency and scalability, enabling deployment on a broad range of healthcare machines. Specifically, I have focused on two main themes: (1) Enabling fusion imaging through ultrasound (US) and Computed Tomography (CT) segmentation, to enhance diagnosis and create a higher-fidelity modality for ablations in operating rooms. (2) Analyzing biomedical signals, particularly electrocardiograms (ECGs), to diagnose CVDs and develop a metric for cardiovascular well-being for data- and resource-constrained environments, including mobile devices. In future, I will advance my vision by focusing on multi-modal solutions that integrate various forms of medical data, while ensuring that these systems remain accessible for resource-constrained settings.

1. Enabling Fusion Imaging: Real-time and Lightweight Networks for Segmentation

At an early stage in my research, I realized the primary challenge hindering the progress of my research vision in medical imaging is the gap in understanding between AI experts and radiologists. Radiologists often lack insights regarding the problems that DL can solve, while AI researchers may not always grasp the key diagnostic challenges in medical imaging modalities. To address this, I conducted a series of surveys to map these knowledge gaps, identify areas where DL can make an immediate impact, and encourage a dialogue that will drive the development of the next generation of efficient and scalable AI-driven diagnostic tools for resource-constrained environments. Initially, I reviewed liver segmentation methods and evaluated their utility in clinical surgeries, highlighting their usage for treatment of hepatocellular carcinoma (HCC or liver cancer) [1]. Simultaneously, I performed risk assessment of computer-aided diagnostic (CAD) software for hepatic resection, quantifying the risks and benefits of adopting CAD systems [2]. Subsequently, I assed efficacy of image fusion strategies immediately post-ablation in liver neoplasms to determine therapeutic response and evaluate the need for follow-up plans [3]. Altogether, these works provide a wholistic understanding to the radiologists regarding the role of DL, CAD, and fusion imaging in clinical surgeries for HCC, highlighting their benefits and risks. I also surveyed advancements in DL architectures for segmentation of B-mode US images, suggesting strategies to improve performance for low-contrast US imaging [4,5]. This review offers valuable insights to AI researchers, highlighting existing challenges that can be addressed through advanced preprocessing techniques, architectural innovations, and the development of novel loss functions.

I laid the foundation of my research vision in medical imaging by enabling fusion imaging, specifically combining real-time US and preoperative CT scans to enhance surgical precision and outcomes. Effective image fusion necessitates accurate segmentation in both US and CT. Consequently, I proposed **Dense-PSP-UNet** [6,7] for real-time US liver segmentation, featuring a dense backbone to minimize parameters and computations. To address poor contrast, I applied contrast-limited adaptive histogram equalization and incorporated pyramid scene parsing for multiscale feature extraction. Finally, I introduced a dynamic loss function combining area- and boundary-based losses for precise liver boundaries. The network achieved excellent performance with 37 FPS, making real-time US-CT fusion viable in the operating room.

The complementary challenge to real-time US segmentation for image fusion is efficient preoperative CT segmentation, especially in developing countries where large workstations may be inaccessible. To address this, I developed **Res-PAC-UNet** [8,9], a lightweight framework for liver CT segmentation featuring a fixed-width residual backbone to minimize computational cost, enabling deployment on a range of devices. I also introduced a pyramid atrous convolution (PAC) module to extract features across CT slices, enhancing segmentation performance. Res-PAC-UNet generates SOTA liver segmentation masks in ≈ 0.5 seconds, using just 15 MB of disk space. Both Dense-PSP-UNet and Res-PAC-UNet have been integrated into an Augmented Reality (AR) medical image visualization toolkit at Hamad Medical Corporation (HMC) for doctors and medical students to explore 3D/fused imaging modalities. My work in biomedical imaging for HCC earned first place in the Concept & Design category at the AI Competition-MENA region 2022 and runner-up at HMC's Star of Excellence Award 2024.

2. Developing Interpretable Metric of Cardiovascular Well-being from ECG

A fundamental aspect of my research vision in biomedical signals has started to take shape through my work on developing a robust metric for cardiovascular well-being using ECG-based age (ECG age) prediction, specifically tailored for data- and resource-constrained environments. Attia et al. [10] demonstrated that neural networks trained to predict age from 12-lead ECG signals inherently capture cardiovascular health, with predicted ECG age closely aligning with chronological age in healthy subjects and deviating in those with CVDs [11,12]. ECG age is valuable for tracking disease progression, with improvements in cardiovascular health reflected by a decrease in ECG age. It is also cost effective and easily interpretable, allowing patients and physicians to monitor the impact of lifestyle changes and treatments.

Two key challenges remain unaddressed in the ECG age literature. First, most studies rely on 12-lead ECGs, which are common in clinical settings but not in emergency services or wearables that typically use 1- to 5-lead monitors (i.e., data constraints). This underscores the need to identify effective lead subsets for ECG age estimation. To address this, I developed a bottom-up search framework to identify high-performing lead subsets, ranging from single- to four-lead configurations. For each subset, I analyzed performance variations and compared them to the standard 12-lead ECG. The insights from the analysis allowed me to stratify the leads and reduce the search space for larger subsets. The results showed that neural networks achieve performance within ~2% of the 12-lead ECG using the proposed four-lead subsets. Importantly, these subsets generalize across neural networks and tasks, including gender estimation and myocardial infarction prediction.

Second, there is a lack of scalable, lightweight, and explainable neural networks for deployment across ECG machines, mobile devices, and smartwatches (i.e., resource constraints). To address this, I developed **ECGNextFormer**, a lightweight hybrid CNN and transformer architecture for efficient ECG age estimation. ECGNextFormer features a shallow backbone for efficient feature extraction, combined with a bottleneck drop key self-attention layer to provide a global receptive field and reweight feature maps. Additionally, a multiscale pooling block reduces temporal features while retaining fine-grained spatial details. These innovations surpass the best networks in literature by 14.7% in performance and 53x in parameter efficiency. ECGNextFormer's interpretability is enhanced with fine-grained heatmaps generated via guided backpropagation, and a novel lead scoring scheme that quantifies ECG lead significance in both healthy and diseased groups.

3. Future Works

My future work aims to advance my research vision by developing *light language models* for medical applications that can seamlessly integrate multi-modal data, combining medical images, biomedical signals, and reports into a unified system. The multi-modal approach will allow for the seamless integration of insights from different data modalities, which is crucial for diagnosing complex diseases with diagnostic clues spread across multiple modalities. In the immediate future, I will focus on enabling multi-modal analysis and zero-shot diagnosis techniques.

Multi-modal analysis is crucial for HCC staging, where accurate diagnosis relies on more than just imaging data. While CT-based segmentation can determine tumor size, count, and location, effective staging also depends on factors like liver function (assessed using Child-Pugh score), extent of cirrhosis, and blood parameters. *light language models* are ideal for aligning information from images, radiological reports, and blood parameters, facilitating accurate staging while supporting widespread deployment.

Zero-shot diagnosis has strong potential for diagnosing rare or unseen CVDs in training data. The key advantage of this approach is that it enables diagnoses without the need for large labeled datasets, making it cost-effective, and time-efficient. A potential design of this approach may involve constructing an expert vector database, combining knowledge from well-established medical diagnostic books. The database can provide expert diagnostic guidance, which can be combined with extracted data features to design prompts for *light language models*. Subsequently, *light language models* can interpret the features under the guidelines of diagnostic guidance and provide a diagnosis.

In my future research, I am committed to mentoring students and actively involving them in my work. For instance, in my research on multimodal analysis that combines imaging, signal, and blood report data, students will assist in processing various data modalities such as CT scans, blood test results, and liver function scores. They will contribute to designing a benchmark dataset for multimodal learning that can be published. For my work on zero-shot diagnosis, students will play a key role in building an expert vector database that is fact-checked, and they will assist in designing diagnostic prompts for *light language models*. To prepare myself for this role, I serve as a mentor for Carnegie Mellon University's Qatar Student Initiated Undergraduate Research Program (QSIURP).

References:

1. **Mohammed Yusuf Ansari**, Alhusain Abdalla, Mohammed Yaqoob Ansari, Mohammed Ishaq Ansari, Byanne Malluhi, Snigdha Mohanty, Subhashree Mishra, Sudhansu Sekhar Singh, Julien Abinahed, Abdulla Al-Ansari, Shidin Balakrishnan, Sarada Prasad Dakua. “Practical utility of liver segmentation methods in clinical surgeries and interventions”. BMC Medical Imaging, (2022).
2. Yusuf Akhtar, Sarada Prasad Dakua, Alhusain Abdalla, Omar Mousa Aboumarzouk, **Mohammed Yusuf Ansari**, Julien Abinahed, Mohamed Soliman Mohamed Elakkad, and Abdulla Al-Ansari. “Risk Assessment of Computer-aided Diagnostic Software for Hepatic Resection”. IEEE Transactions on Radiation and Plasma Medical Sciences, (2021).
3. Pragati Rai, **Mohammed Yusuf Ansari**, Mohammed Warfa, Hammad Al-Hamar, Julien Abinahed, Ali Barah, Sarada Prasad Dakua, Shidin Balakrishnan. “Efficacy of fusion imaging for immediate post-ablation assessment of malignant liver neoplasms: A systematic review”. Cancer Medicine, (2023).
4. **Mohammed Yusuf Ansari**, Iffa Afsa Changaai Mangalote, Pramod Kumar Meher, Omar Aboumarzouk, Abdulla Al-Ansari, Osama Halabi, Sarada Prasad Dakua. “Segmentation of B-mode Ultrasound Images Using Neural Networks: A Comprehensive Survey”. IEEE Transactions on Emerging Topics in Computational Intelligence, (2024).
5. **Mohammed Yusuf Ansari**, Marwa Qaraqe, Erchin Serpedin, Raffaella Righetti, Khalid Qaraqe. “Unveiling the Future of Breast Cancer Assessment: A Critical Review on Generative Adversarial Networks in Elastography Ultrasound”. Frontiers in Oncology, (2023).
6. **Mohammed Yusuf Ansari**, Pragati Rai, Yin Yang, Shidin Balakrishnan, Julien Abinahed, Abdulla Al-Ansari, Omran Almokdad, Ali Barah, Ahmed Omer, Ajay Vikram Singh, Pramod Kumar Meher, Jolly Bhadra, Osama Halabi, Mohammad Farid Azampour, Nassir Navab, Thomas Wendler, Sarada Prasad Dakua. “Dense-PSP-UNet: A neural network for fast inference liver ultrasound segmentation”. Computers in Biology and Medicine, (2023).
7. **Mohammed Yusuf Ansari**, Iffa Afsa Changaai Mangalote, Dima Masri, Sarada Prasad Dakua. “Neural Network-Based Fast Liver Ultrasound Image Segmentation”. IEEE international joint conference on neural networks (IJCNN), 2023.
8. **Mohammed Yusuf Ansari**, Yin Yang, Shidin Balakrishnan, Julien Abinahed, Abdulla Al-Ansari, Mohamed Warfa, Omran Almokdad, Ali Barah, Ahmed Omer, Ajay Vikram Singh, Pramod Kumar Meher, Jolly Bhadra, OsamaHalabi, Mohammad Farid Azampour, Nassir Navab, Thomas Wendler, and Sarada Prasad Dakua. “A lightweight Neural Network with Multiscale Feature Enhancement for Liver CT Segmentation”. Scientific Reports, (2022).
9. **Mohammed Yusuf Ansari**, Snigdha Mohanty, Serah Jessy Mathew, Subhashree Mishra, Sudhansu Sekhar Singh, Julien Abinahed, Abdulla Al-Ansari, and Sarada Prasad Dakua. “Towards Developing a Lightweight Neural Network for Liver CT Segmentation”. International Conference on Medical Imaging and Computer-Aided Diagnosis, (2022).
10. Zachi I. Attia, Paul A. Friedman, Peter A. Noseworthy, Francisco Lopez-Jimenez, Dorothy J. Ladewig, Gaurav Satam, Patricia A. Pellikka, Thomas M. Munger, Samuel J. Asirvatham, Christopher G. Scott, Rickey E. Carter, and Suraj Kapa. “Age and Sex Estimation Using Artificial Intelligence from Standard 12-Lead ECGs”. Circulation: Arrhythmia and Electrophysiology, (2019).
11. **Mohammed Yusuf Ansari**, Marwa Qaraqe, Fatme Charafeddine, Erchin Serpedin, Raffaella Righetti, Khalid Qaraqe. “Estimating Age and Gender from Electrocardiogram Signals: A Comprehensive Review of the Past Decade”. Artificial Intelligence in Medicine, (2023).
12. **Mohammed Yusuf Ansari**, Marwa Qaraqe, Erchin Serpedin, Raffaella Righetti, Khalid Qaraqe. “Enhancing ECG-Based Heart Age: Impact of Acquisition Parameters and Generalization Strategies for Varying Signal Morphologies and Corruptions”. Frontiers in Cardiovascular Medicine, (2024).