Hanshi Sun

Personal Statement

The COVID-19 pandemic induced a significant strain between the demand for and relatively fixed supply of available physician labor, thereby threatening its ability to function properly in the future. As such, and in memory of my grandfather who died due to insufficient medical resources, I resolved to develop intelligent healthcare solutions to help more people avoid the misery of disease and premature death. Witnessing the development of artificial intelligence has reinforced my belief that machine learning can help medical professionals better use their expertise to make decisions and mitigate the labor shortage, especially for medical imaging for screenings, precision medicine, and risk assessment. However, the community has had difficulty building unified diagnosis models for several common diseases. Therefore, I am pursuing the M.S. in Computer Science at Columbia University where I will refine the skills needed to contribute to machine learning and medical imaging by developing a multi-modal multi-task unified model for diagnosing multiple diseases.

My academic trajectory is like a gradient descent algorithm $\theta_{n+1} = \theta_n - \eta \nabla F(\theta_n)$, where θ is my life goal, η is the step length, $\nabla F(\theta)$ is my research interest (descent direction), and the most important thing is my resolution to keep moving to the next point in the curve. In this regard, my undergraduate studies in Electronic Science and Technology at Southeast University exposed me to several promising machine learning projects for medical diagnosis, including medical imaging and signal processing. For example, under the supervision of Professor Hao Liu at the National ASIC Engineering Research Center, I led the "Arrhythmia Detection" project which aimed to design and deploy an efficient arrhythmia detection algorithm on resource-constrained devices. After noticing quantization methods suffered from accuracy degradation, I suggested pruning useless parameters and emphasizing essential network layers. Inspired by existing methods that quantize weights and activations of deep neural networks into multiple binary bases, I published two papers that proposed a 1D adaptive loss-aware multi-bit networks quantization method to allocate different bitwidth for respective layers, balancing high performance (95.84% accuracy) and low resource consumption (23.36× compression rate). This experience confirmed the prospect of deployment on wearable devices to realize a real-time arrhythmia diagnosis, prompting me to investigate the difficulty of deploying more sophisticated models to support 2D medical images.

To extend my domain expertise in this area, I joined the PAttern Learning and Mining (PALM) Lab to focus on medical imaging, especially fundus and chest X-ray images. Previous models mainly focused on a specific task and modality, meaning the information was restricted from a single source. To address this issue, our team set out to develop a multi-modal multi-task model for massive datasets from different disease domains. Initially, I found the model suffered greatly from medical label uncertainty and inconsistency. Taking chest X-rays as an example, most datasets have employed rule-based natural language processing (NLP) to radiology reports to extract labels automatically which frequently results in label noise. Therefore, I introduced inter- and intra-instance consistency regularization to provide consistent predictions for images with similar medical findings to combat noisy labels. Next, I trained models on multi-datasets with data imbalance to take advantage of the attention mechanism and correlation between tasks. One finding showed that a unified model trained under such settings could outperform single-task models even when downstream task domains lacked similarity; however, executing the development still posed challenges.

Leveraging my exposure to medical imaging, I joined the University of Alberta as a Research Intern where I was tasked with designing a model lacking prior abnormality information to identify and localize anomalies in optical coherence tomography (OCT) images. Initially, I wanted to build a hybrid architecture that incorporated many models to improve detection performance as it's the easiest way to simultaneously balance respective outputs. However, my professor proceeded to exclaim: "*What is the motivation of the method? Research is not piling others' ideas and pursuing model accuracy without a big picture. Hanshi, you should think critically about whether it makes sense and keep striving to become a scientist*." After comprehensively investigating pathological features, I discovered the uniqueness of the OCT images, different from natural images, have notable regular structure features (e.g., structured anatomy). Stimulated by this finding, I presented an architecture to integrate segmentation knowledge as privileged features, incorporating privileged features distillation and reverse distillation methods with the help of structure extraction networks. The result far exceeded my expectations, outperforming state-of-the-art (SOTA) methods such as STFPM by 5.76%. Ultimately, I realized that medical images have high similarity in overall structure and diversity in details; however, encoding the uniqueness of medical images precisely into the model is a challenging task.

Currently, I am working as an R&D iPad System Intern at Apple. One of my assignments is to capture and reduce the camera noise for coexistence testing. My work and research experiences have shown it is typical to find noise in images. For some use cases, noise is not disruptive; however, for medical imaging, one small noise spot can significantly impact a diagnosis. Conscious of this issue, my time at Apple has empowered me to define my future research direction and gain a sense of image noise problems which can be challenging to research. Presently, the medical imaging community mainly concentrates on a single task rather than a unified generalized model. Considering the patient privacy-preserving and the availability of datasets, I plan to build a multi-modal multi-task continual learning model for diagnosing multiple diseases. With the help of a comprehensive model, doctors can make judgments more efficiently, thereby making medical diagnoses available to a greater number of people.

I am drawn to the program due to Columbia's leading position in machine learning and computer vision research. It is my dream to join Professor Richard Zemel's group, whose research interests in few-shot learning and continual learning align perfectly with my focus. I admire Professor Richard's recent work, "Learning a Universal Template for Few-shot Dataset Generalization," which demonstrated a partially-parameterized model to readily specify a good feature extractor for a held-out dataset. As part of this group, I aspire to explore similar approaches to reduce the annotation burden and build the aforementioned unified generalized diagnosis model. I am also interested in Professor Carl Vondrick's group and their projects on multi-task learning, which is a promising way to predict multiple diseases. Moreover, Professor Elias Bareinboim's projects on causal inference and Professor Ansaf Salleb-Aouissi's projects on semi-supervised machine learning are intriguing. Beyond research, Columbia's customizable curriculum will enable me to acquire the knowledge needed to build my ideal diagnosis model, including security and networking, given the need to protect patients' private data. Upon completing the program, I plan to pursue a Ph.D. before progressing to a career in R&D or academia. I look forward to utilizing my time as a Columbia graduate student to explore the diverse application areas of my specialized knowledge and look forward to taking this next step in my professional development.