Recent advancements in machine learning have shown great potential for medicine. However, few algorithms are regularly used in clinical care due to several key challenges: (1) treating patients involves critical decisions, and the models involved must be carefully evaluated in clinical trials; (2) integrating models into clinical workflows is a complicated and intricate process; and (3) the complexity of real-world data makes it difficult to train robust and reliable models.

My research aims to address these challenges: I build **novel machine learning algorithms** for **clinically and scientifically important problems**, working closely with clinicians to build tools to **integrate and evaluate these algorithms in clinical workflows**, and I identify and address the underlying **difficulties of using real-world data** not apparent in benchmark datasets.

My work builds methods for analyzing and understanding patient health at a wide range of scales, ranging from the patient level in cardiology, to the cellular level in pathology, and to the molecular level in RNA sequencing. Many real-world challenges are shared across these scales, including limited and noisy labels for the data, complex interactions between data points (e.g. the temporal relationship between samples from the same patient and the 3D structure underlying 2D images of the same organ), and intricate relationships between data from a mixture of modalities.

Looking forward, I plan to continue building strong interdisciplinary connections as a professor. My research group will develop and deploy new algorithms for problems across medicine and other areas of science and work to identify and address the shared real-world challenges.

Prior work

My work spans several broad areas, which I group by theme and discuss in the following sections.

Video-based deep learning for assessing heart function One of the most widely used diagnostic tools in cardiology is the echocardiogram, an ultrasound of the heart. To properly diagnose patients and track changes in health over time, numerous measurements (e.g. the amount of blood pumped per heartbeat or the thickness of a chamber wall) are taken across studies consisting of several dozen videos from a broad array of views. These measurements are a tedious but critical process currently performed by sonographers and cardiologists, requiring the tracking of heart movement over each beat.

To help streamline the clinical workflow, I developed EchoNet-Dynamic [1] to automate the evaluation of the ejection fraction of the heart, a key measurement regarding how efficiently the heart is pumping blood. EchoNet-Dynamic uses video-based deep learning to provide an estimate of the ejection fraction from a raw echocardiogram video, as well as segmentations of the left ventricle, the chamber responsible for pumping blood throughout the body, to provide feedback to clinicians. The predictions by EchoNet-Dynamic reached human-level accuracy, staying well within the range of typical measurement variation between different clinicians.

The automated workflow enabled by EchoNet-Dynamic additionally allows clinicians to assess echocardiograms in several ways that were previously too labor-intensive to regularly use. First, the function of the heart can have considerable variation between beats, particularly in patients

with poor cardiac function, and automated measurements allow rapid assessment of the beat-to-beat variability of a patient. Second, the current clinical workflow typically focuses on two key frames of a heartbeat (diastole and systole), while the automated segmentations allow the analysis of heart motion frame-by-frame, with temporal resolutions as low as 20 milliseconds. The methods developed for EchoNet-Dynamic are also more broadly applicable for other uses in interpreting echocardiograms, including other diseases and biomarkers [2, 3, 4].

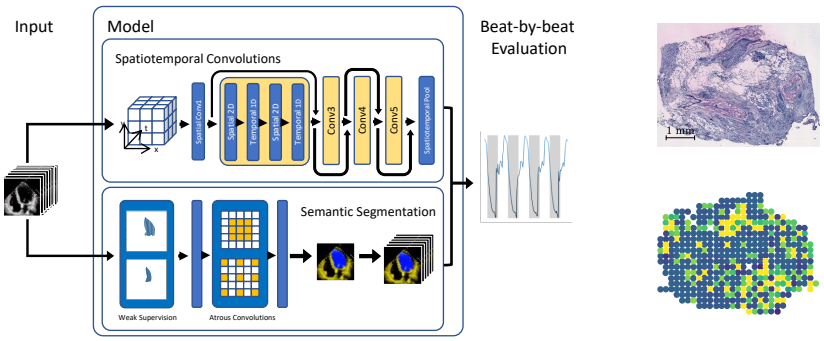
Deployment and evaluation of machine learning in clinical workflows Integrating machine learning into clinical workflows is essential for leveraging its full capabilities. However, this process is challenging: the existing workflow cannot be disrupted, the predictions must be provided with low latency, and the system must be secure due to the sensitivity of healthcare data and privacy concerns. This integration is also necessary for studying and understanding the downstream interactions between the model's predictions and clinical care.

In light of this, I built the tools and infrastructure needed to integrate EchoNet-Dynamic into the standard clinical workflow. This allowed us to conduct a blinded, randomized clinical trial consisting of 3,769 echocardiogram studies to compare the model's performance against sonographer assessment [5]. First, the trial found that the model's predictions were well integrated into the workflow, and cardiologists reviewing the initial assessments could not reliably determine if the model or a sonographer made the initial assessment. The trial also found that cardiologists made substantial changes less frequently to the model's initial assessments (16.8% versus the sonographer's 27.2%) and that using the model's initial assessment saved both cardiologist and sonographer time. The model is currently late in the FDA 510(k) clearance process to ensure safety and efficacy.

In addition to the real-world impact, my work on heart ultrasounds has also addressed two real-world instances of domain shift, a key challenge for machine learning. First, I developed methods for adapting methodology for adult patients to pediatric patients [6]. Due to the rarity of pediatric heart disease, less training data is available. Pediatric patients also have increased abnormalities, heart rates, and size variability. My work focused on using transfer learning to share information from the adult EchoNet-Dynamic dataset to the pediatric setting. Second, I worked on developing methods for point-of-care ultrasounds in the emergency department [7]. Patients in the emergency department are often less stable, resulting in much lower-quality videos that are challenging to assess. My work on emergency department ultrasounds added the ability to evaluate video quality to the models, allowing the high-quality videos from each patient to be identified and used for further assessment.

Linking cell morphology to gene and protein expression The spatial organization of gene and protein expression has important biological effects on the properties of the tissue. An important question is how to link the microscopy images of a tissue with the gene and protein expression, allowing us to better understand how cell morphology relates to the function of the tissue.

My work developing ST-Net studies the relationship between cell morphology and spatial transcriptomics in breast cancer and can handle the gigapixel scales typical of microscopy images [8, 9]. Spatial transcriptomics allows for the direct measurement of RNA abundance with high spatial



(a) (b)

(a) EchoNet-Dynamic provides clinicians with an assessment of echocardiogram videos.
(b) ST-Net links the morphology of cells with gene expression.

resolution, allowing us to link the morphology of cells to the local gene expression. ST-Net was able to identify 102 genes whose expression levels could be identified from cell morphology alone. These genes were involved in known cancer, pharmacogenomic, and immune pathways, indicating that these pathways have biological effects visible in the morphology of tissue.

I have also developed methods to identify Alzheimer's disease lesions from gigapixel-scale microscopy images [10] and to predict patient phenotypes from single-cell RNA-sequencing data [11].

Learning representations across modalities without labeled data In medicine, different imaging modalities and diagnostic tools are used to gather information about a patient from different perspectives. Experts are needed to generate reliable labels, resulting in difficulties in generating large labeled datasets. In the most limited setting, there are no explicit labels for training, and we would only know which data samples were obtained from each patient. With no labels for the data, it can be challenging to learn a meaningful representation linking the modalities together.

In an ongoing project, I am developing methods to learn representations connecting data from four diagnostic modalities: echocardiograms, electrocardiograms (EKGs), angiograms, and chest X-rays. A key insight is that although data from these modalities differ significantly in appearance, samples from the same patient are highly related and implicitly carry a large amount of clinical information. For example, cardiologists can identify the heart region with damage from a heart attack, along with the severity, from both echocardiograms and EKGs. As a result, paired samples from the same patient, even from different modalities, contain large amounts of implicit labeling.

Taking advantage of this insight, my work uses contrastive learning to bring the representation of samples from the same patient closer together, while pushing samples from different patients further apart. The resulting representations can retrieve paired samples from the same patient across modalities, ranking the paired sample at over the 95th percentile on average. The representations also allow cheaper and faster modalities to be used to estimate measurements from the modalities that are more difficult to obtain.

Research Vision

There are many opportunities to impact medicine and other areas of science with machine learning. As shown by my prior work, I have extensive experience developing machine learning methods to leverage the potential of these opportunities. I discuss several of the areas I will work on, along with the challenges involved in these applications.

Training generalist models with specialized knowledge A common paradigm in machine learning is to train a model on a large, general dataset, followed by fine-tuning it on a smaller but much more targeted and specific dataset. This method often improves the performance on the second task but results in the model no longer being able to perform well on the first task.

An application of this training technique is to improve model performance on rare diseases, where the datasets are typically small due to the rarity of the disease. In these settings, losing the ability to perform on the original, more general dataset, is highly problematic, potentially resulting in a model that performs well on the rare disease, while performing poorly on standard cases.

Understanding interconnected data Data in medicine is often tightly interconnected. In echocardiogram studies, individual videos are 2D slices from a 3D structure. In pathology, multiple stains are used to highlight different proteins and structures of interest from the same tissue. Across medicine, samples from the same patient are taken across time. Current models are often trained on these data points independently, without understanding the relationship between the samples.

This lack of understanding can degrade model performance and prevent models from gauging their confidence. A downstream impact of this inability is that models sometimes can output a best guess when the provided information is insufficient to yield an answer with any confidence. This can be highly problematic, as the correct behavior would be to ask for more information or request additional tests, rather than making a potentially inaccurate guess.

Building medical foundation models Building models to learn the multimodal interactions between text and images is an area of great interest. Promising results have already been shown for natural images, but medical data is often far more complex. In many medical applications, reports refer to particular areas of interest in the image. For example, pathology reports often refer to a tiny and specific region of a gigapixel image, and echocardiogram reports may identify abnormalities only visible in a small subset of videos. Learning the interactions between the text and images in these settings is challenging for existing techniques.

More broadly, I aim to build an interdisciplinary and collaborative research agenda utilizing both machine learning and domain expertise. My research vision is to utilize these combined skillsets to build and deploy tools with real impact in medicine and other areas of science.

References

[1] David Ouyang, Bryan He, Amirata Ghorbani, Neal Yuan, Joseph Ebinger, Curtis P Langlotz, Paul A Heidenreich, Robert A Harrington, David H Liang, Euan A Ashley, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*, 2020.

[2] Grant Duffy, Paul P Cheng, Neal Yuan, Bryan He, Alan C Kwan, Matthew J Shun-Shin, Kevin M Alexander, Joseph Ebinger, Matthew P Lungren, Florian Rader, et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiology*, 2022.

[3] Grant Duffy, Shoa L Clarke, Matthew Christensen, Bryan He, Neal Yuan, Susan Cheng, and David Ouyang. Confounders mediate AI prediction of demographics in medical imaging. *npj Digital Medicine*, 2022.

[4] J Weston Hughes, Neal Yuan, Bryan He, Jiahong Ouyang, Joseph Ebinger, Patrick Botting, Jasper Lee, John Theurer, James E Tooley, Koen Nieman, et al. Deep learning evaluation of biomarkers from echocardiogram videos. *EBioMedicine*, 2021.

[5] Bryan He, Alan C Kwan, Jae Hyung Cho, Neal Yuan, Charles Pollick, Takahiro Shiota, Joseph Ebinger, Natalie A Bello, Janet Wei, Kiranbir Josan, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. *Nature*, 2023.

[6] Charitha D Reddy, Leo Lopez, David Ouyang, James Y Zou, and Bryan He. Video-based deep learning for automated assessment of left ventricular ejection fraction in pediatric patients. *Journal of the American Society of Echocardiography*, 2023.

[7] Bryan He, Dev Dash, Youyou Duanmu, Ting Xu Tan, David Ouyang, and James Zou. AI-enabled assessment of cardiac function and video quality in emergency department point-of-care echocardiograms. *The Journal of Emergency Medicine*, 2023.

[8] Bryan He, Ludvig Bergenstråhle, Linnea Stenbeck, Abubakar Abid, Alma Andersson, Åke Borg, Jonas Maaskola, Joakim Lundeberg, and James Zou. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nature Biomedical Engineering*, 2020.

[9] Ludvig Bergenstråhle, Bryan He, Joseph Bergenstråhle, Xesús Abalo, Reza Mirzazadeh, Kim Thrane, Andrew L Ji, Alma Andersson, Ludvig Larsson, Nathalie Stakenborg, et al. Super-resolved spatial transcriptomics by deep data fusion. *Nature Biotechnology*, 2022.

[10] Bryan He, Syed Bukhari, Edward Fox, Abubakar Abid, Jeanne Shen, Claudia Kawas, Maria Corrada, Thomas Montine, and James Zou. AI-enabled in silico immunohistochemical characterization for Alzheimer's disease. *Cell Reports Methods*, 2022.

[11] Bryan He, Matthew Thomson, Meena Subramaniam, Richard Perez, Chun Jimmie Ye, and James Zou. CloudPred: Predicting patient phenotypes from single-cell RNA-seq. In *Pacific Symposium on Biocomputing*, 2022.