

Dissecting the potential for AI in pathology

November 2021—*The following is an edited excerpt of the article “Attention-based deep multiple instance learning,” written by Jonathan Glaser, a recent graduate of the computer science and biotechnology master of science degree programs at New York University Tandon School of Engineering, in Brooklyn. The excerpt delves into how aspects of artificial intelligence can transform health care, and pathology in particular. To read the full article, go to <https://tinyurl.com/AI-based-learning>.*

Much of AI’s momentum today can be attributed to the success of deep neural networks, which would not be possible without a perfect storm of the following four driving forces:

1. The increasing availability of massive data sets, such as ImageNet’s 15 million labeled images; Facebook’s library of billions of images; YouTube’s video library, which grows by 300 hours of video per minute; and Tesla’s collection of driving data, which adds 1 million miles of data per hour.¹
2. The use of graphic processing units (GPUs) and, later, more AI-specialized hardware called tensor processing units (TPUs), which are optimized for training deep-learning models. TPUs consist of many cores, which enable them to process large amounts of data and perform multiple computations in parallel. A 2018 report by OpenAI proposed that prior to 2012, the growth of computing power closely tracked Moore’s law, doubling every two years, and that post-2012, computing power has been doubling every three to four months. Overall, since 2012, this metric has grown by a factor of more than 300,000, while a two-year doubling period would have only yielded a 16x increase.²
3. The availability of cloud computing, which has made storage of large data sets and use of those data sets in training models more accessible and economical.
4. Open-source algorithmic development modules, such as Facebook’s PyTorch, Google’s TensorFlow, Microsoft’s Cognitive Kit, and others.

As Eric Topol, MD, founder and director of the Scripps Research Translational Institute, among others, asserts, the rise of AI offers an exciting opportunity to revolutionize the health care industry. However, the extent to which health care becomes infused with AI needs to be tempered by the industry’s inherent need for empathy and connection between clinicians and patients. A physician affords patients a sense of ethics and core values, neither of which can be replicated by a computer. Assertions that AI will soon become sophisticated enough to lead to full automation in other industries has stoked fears among medical professionals about whether human involvement in medicine will become a thing of the past. However, the possibility of AI leading to full automation in medicine is still far off. Those involved in AI have time to strike the right balance between doctors and machines.

Keeping a human in the loop warrants the need for transparency into how an algorithm arrived at a diagnosis, recommendation, or prediction. Explainable AI is a set of processes or methods that provides this insight and is crucial for building trust among those relying on AI-enabled systems while also ensuring accuracy, fairness, and compliance with regulatory standards. Explainability provides clinicians with quality control and checks and balances, and it can help them become more confident about relying on algorithms to make final diagnoses.

Robust models will need to be developed with a satisfactory level of explainability. But if done correctly, the resulting increased workflow and efficiency can afford clinicians more time to connect with patients. Paradoxically, the rise of machines can restore the humanity in medicine and allow medical professionals to get back in touch with their motivations for pursuing a medical career in the first place.

Artificial intelligence in pathology

One of the most effective applications of AI in health care has been in medical imaging. Radiology, pathology, and dermatology are specialties that rely on visual pattern analysis and are, therefore, positioned to undergo a rapid and dramatic transformation due to integration with AI. Here, we focus on what that might look like for pathology.

The practice of using microscopes to examine glass slides containing tissue samples has been largely unchanged for a century. In recent years, however, slides increasingly are being digitized using digital slide scanners to produce whole slide images that can be examined on a computer. Yet pathologists have been slow to adopt WSI and other digital techniques, which has resulted in the encroachment of AI into pathology being slower than expected. Nevertheless, WSIs have laid the groundwork for incorporating neural network image processing into pathology, thereby making a new AI-assisted era in the field imminent.¹

One set of AI technologies has been directed towards simplifying routine workflows that are typically performed by pathologists—for example, detection of tumor tissue in biopsy samples and determination of tumor subtype based on morphology. A major milestone for AI in pathology was the Camelyon16 challenge, which set the goal of developing algorithms to detect metastatic breast cancer in WSIs of lymph node biopsies. The data set provided was made up of 400 WSIs in which pathologists manually delineated regions of metastatic cancer and is one of the largest labeled pathology data sets. This enabled the team with the top submission, whose algorithm performed on par with pathologists, to make use of supervised learning.³

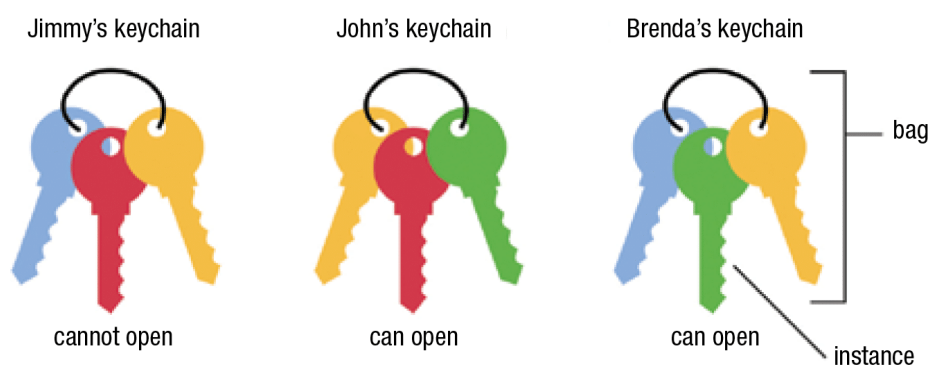
In general, supervised learning is a machine-learning approach in which an algorithm is shown a sequence of input data and corresponding output labels to detect the underlying pattern that reveals the relationship between the inputs and outputs. This technique allows the algorithm to accurately label data that it hasn't seen before and can be used for classification or regression tasks.

A major downside of supervised learning is that it typically requires the training data set to be hand-labeled by domain experts. This is particularly the case when dealing with WSIs. For these and other reasons, designing supervised learning models for everyday use in pathology is highly impractical.⁴

Whole slide image classification using deep multiple instance learning

Multiple instance learning (MIL) is a variation of supervised learning that is more suitable to pathology applications. It involves assigning a single class label to a collection of inputs—in this context, referred to as a bag of instances. This formulation allows us to leverage weakly labeled data and naturally fits various problems in such diverse fields as computer vision and document classification.

A simple example is shown in the figure below. In the example, we only know whether a keychain contains the key that can open a given door. We do not have any information about the individual keys. This allows us to infer that the green key can open the door. Similarly, in medical imaging applications, an overall patient diagnosis can be a substitute for costly local annotations provided by an expert.



The MIL formulation is appropriate for computer-aided diagnoses for four main reasons:

1. Processing WSIs consisting of billions of pixels is computationally infeasible. For scale, roughly 470 pathology

images contain approximately the same amount of pixels as the entire ImageNet data set.⁴ Therefore, it is desirable to divide each medical image into smaller tiles that can be further considered a bag with a single label.

2. Supervised approaches exist that require pixel-level annotations that outline abnormalities in a patient's medical scan. However, these require pathologists to spend large amounts of time on data preparation, thereby interfering with their daily routines. Since MIL techniques only require weak labels (i.e. the overall diagnosis of a patient), they hold the promise of significantly reducing pathologist workloads.

3. MIL naturally fits the task of imaging-based patient diagnosis. Diseased tissue samples have both abnormal and healthy regions, while healthy tissue samples have only healthy regions. These can easily be represented as bags, each with a single label.

4. It is possible to formulate MIL as a deep-learning problem. This allows for the creation of a smooth end-to-end pipeline for pathology in which WSIs are fed in as the input and diagnoses are returned as the output. Deep MIL models can also be made to be highly interpretable, which caters to the explainability requirement of AI systems in health care discussed earlier.⁵

1. Topol EJ. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books; 2019.
2. Amodei D, Hernandez D. AI and compute. OpenAI; May 16, 2018. <https://openai.com/blog/ai-and-compute>
3. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. arXiv; June 18, 2016. <https://arxiv.org/abs/1606.05718>
4. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301–1309.
5. Ilse M, Tomczak JM, Welling M. Attention-based deep multiple instance learning. arXiv; June 28, 2018. <https://arxiv.org/abs/1802.04712>

Mount Sinai establishes AI and human health department

The Icahn School of Medicine at Mount Sinai has launched the department of artificial intelligence and human health, with the goal of further advancing artificial intelligence to transform health care.

"The overarching goal of the department of AI and human health is to impact patients' health with AI," said Thomas J. Fuchs, Dsc, dean of artificial intelligence and human health at Icahn Mount Sinai, New York, in a press statement. "We will accomplish this," he explained, "by building AI systems at scale from data representing Mount Sinai's diverse patient population. These systems will work seamlessly across all hospitals and care units to support physicians, foster research, and, most importantly, help patient care and well-being."

Dr. Fuchs and his team will work on integrating machine learning and AI-driven decision-making throughout the health system's eight hospitals. This will include creating a hub-and-satellite model to make new tools and techniques available to all Mount Sinai physicians and building an infrastructure for high-performance computing and data access to improve the organization's diagnostic and treatment capabilities.

The department will be housed in a new research facility on campus that is expected to open in late 2022.

Earlier this year, Icahn Mount Sinai announced it will offer a PhD concentration in artificial intelligence and

emerging technologies in medicine as part of its PhD in biomedical sciences program, starting in the fall of 2022.

PreciseMDX releases digital shipping kits for diagnostic tests

PreciseMDX is offering digital shipping kits for at-home and on-site diagnostic tests to allow laboratories, patients, and providers to track the entire testing process and, thereby, increase efficiency.

Via the digitized process, patients scan a QR code on the shipping kit, which connects them to the laboratory and prescribing provider so that all parties are notified that testing has begun. The end-to-end solution also informs the patient that the lab received the specimen, initiating the insurance billing process, and notifies the provider and patient of the test results.

The shipping kits support a variety of diagnostic tests. PreciseMDX also provides multichannel distribution options so patients can obtain test kits in the provider office, over the counter, or via mail, vending machines, or collection and drop-off sites.

[PreciseMDX](#), 866-462-4677

Dr. Aller practices clinical informatics in Southern California. He can be reached at raller@usc.edu. Dennis Winsten is founder of Dennis Winsten & Associates, Healthcare Systems Consultants. He can be reached at dwinsten.az@gmail.com.