LOOKING INTO THE FUTURE

PATHOLOGY AND ARTIFICIAL INTELLIGENCE: A DREAM TEAM?

Yfke Prins[1]

[1]Master's student medicine, Radboud university medical center, Nijmegen, the Netherlands

Insight

The field of pathology encompasses the understanding and visualisation of illnesses on a cellular level, allowing for the specific characterisation of disease. For this reason, pathology is essential in clinical care. Nevertheless, despite its necessity, only a minority of the students choose this profession. The field of pathology will change in the future, mostly because of the exciting opportunities for the application of artificial intelligence in pathology. It is clear that this shift is approaching fast. However, when and to what extent remains undefined. Will artificial intelligence have received a prominent place in the doctor's office by 2050, or will the translation to the clinic prove to be more difficult?

Pathology is –arguably– the backbone of the diagnostic process, as the examination of histological slides often provides the highest level of certainty about a diagnosis. Especially in oncology, determining the type of malignancy, often including genetic testing of the tissue, is indispensable. Up until five years ago, the histological slides were predominantly examined through the microscope by pathologists. However, in the past five years, pathology has started to transition from microscopy to digital pathology, meaning that the histological slides are scanned and stored on the computer [1]. Pathologists can then examine the slides digitally and even share them online with colleagues. The online images of the histological slides are referred to as Whole Slide Images (WSI) [1]. The implementation of these WSI has opened the door to applying artificial intelligence (AI) in pathology to aid the examination of these slides [1]. The possible uses of AI are auspicious: however, it also raises questions about its future applications. How impactful will AI eventually be in the diagnostic process? Will we receive our diagnosis from an algorithm, a pathologist, or both?

The history of deep learning

Artificial intelligence refers to any intelligence or cognition that a machine displays, similar to the intelligence of animals, such as humans [2]. There are multiple forms of AI, one of which is deep learning [2, 3]. Deep learning itself has emerged in the past ten years and has been applied extensively in other disciplines, such as radiology, speech and image recognition, and pathology. An important enabler for deep learning in pathology is the widespread implementation of digital pathology [1]. One of the barriers to deep learning research, however, is the need for large annotated datasets of WSI to adequately train the algorithms [1, 2]. Nevertheless, in the past few years, more and more datasets have become publicly available, increasing the accuracy of the algorithms, thereby increasing the applications of AI. The computational pathology group of the Radboudumc has been at the forefront of sharing annotated datasets publically, *e.g.* through the CAMELYON challenge in 2016 and 2017, where the research group put a dataset of WSI containing annotated metastases of breast cancer in lymph nodes online [1, 4]. The contest was eventually won by Harvard University and MIT out of 32 participants [1]. During the CAMELYON challenge, the group of Harvard already designed algorithms that could outperform the pathologist used in the challenge. Since then, more and more algorithms have been designed that performed equally to experienced pathologists – and in some cases, even outperformed pathologists [1]. In other words, the AI algorithms enabled better image analysis in certain tasks compared to a trained pathologist [1, 4]. Should this mean that AI needs to take over the profession of being a pathologist to increase adequate patient management?

Computational pathology and deep learning

In order to discuss the future role of deep learning in pathology, we have to get technical. In short, computational pathology concerns the implementation of deep learning algorithms for analysing histopathological images [2]. Computational pathology uses algorithms to analyse slides. The algorithms are developed in order to achieve a specific output, *e.g.* detection of metastases in lymph nodes (Figure 1). The algorithms are trained using datasets of WSI, which are annotated by professionals (mostly pathologists), defining the desired output of the model [2]. If the algorithm is designed to detect metastases in lymph nodes, for example, the input will be WSI, in which the area with metastases will be annotated. In most cases, training will include cases with normal tissue as a control group.

The deep learning algorithm, which uses a neuronal network, operates directly on the WSI pixels, in order to analyse its input. The neuronal network consists of layers of connected neurons, in which, starting from the pixel values, activations are calculated for each consecutive layer on the basis of the values in the previous layer(s). (Figure 1). You can envision this as the connections in our brain – our neurons receive input (the pixels) and connect with thousands of other neurons. In the algorithm, it is not just the value of one pixel that is important but also the connections between the pixels [2]. The nucleus/cytoplasm ratio, for example, is of importance when detecting cancer and can be determined by differences in the intensity of the HE-stain between different parts of the cell; which includes multiple pixels. The algorithm analyses these connections between the pixels in the neuronal network and assigns to each connection a value depending upon the importance of the connection for the outcome of the algorithm. This is a step-wise process, where the algorithm

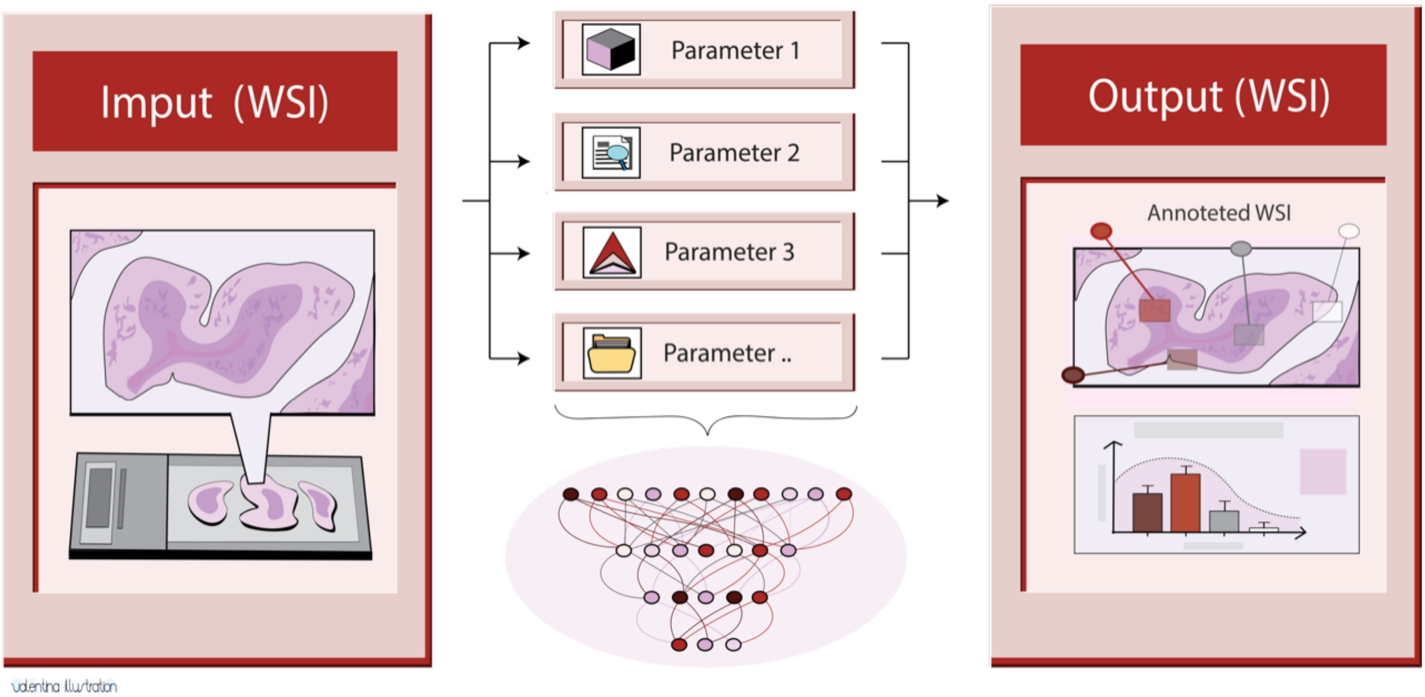

Figure 1: *The algorithm identifies areas of interest (depending upon the algorithm's purpose), through analysis of the connections between pixels. This analysis is achieved by applying a neuronal network to the image, where different layers of the neurons analyse the connections. This is a step-wise process where the value (importance) of previous connections between pixels and/or other connections in which the pixels are involved are also taken into account. Thus, the algorithm narrows down the areas of interest step-by-step, eventually leading to the desired output.*

narrows down the number of connections to those of importance [2]. Thus, the algorithm determines a set of parameters that are essential for establishing the output. Those parameters can then be applied to new WSI after proper training of the algorithm. The algorithm's performance is checked by running it on a second, independent data set and comparing its outcome with the performance of experienced pathologists.

The advantages of deep learning

One of the most significant advantages of applying deep learning is that AI is consistent in its analysis of the slides, minimising inter-observer differences. The algorithm functions the same regardless of which slides it analyses and should consistently provide the correct output. However, if a pathologist analyses the slides, you will always have a certain degree of disagreement and inter-observer differences [1, 2, 4]. Furthermore, the algorithms can analyse data faster than a pathologist [1-4]. Thus, its application comes in handy when there is ever-more tissue that needs to be analysed. For example, deep learning could screen large amounts of slides, which would be useful in nationwide screening programs. These programs, such as the screening programs for colon cancer, and cervical cancer, often generate many tissue samples that take quite some time to analyse and where the predominant part will not contain malignancies [1]. Conversely, the algorithm could mark any WSI where malignancy is suspected, which a pathologist can then analyse, saving a lot of time [1, 2].

Furthermore, deep learning can be used for quantitative tasks, *e.g.* determining the amount of tumour-infiltrating lymphocytes, mitosis hotspots, or programmed death-ligand 1 expression, therefore being able to also determine a wide range of biomarkers that are too work-intensive to be determined by the pathologist [4]. By obtaining more information about these biomarkers, we might even be able to better correlate these biomarkers to clinical outcomes and include more biomarkers in our prognosis or treatment choice. In addition, deep learning can pick up more subtle changes in an image, those which human eyes often overlook. An example of this is tumour budding, which refers to a single or a few cancerous cells at the edges of a tumour, indicating a more invasive character [1, 4]. Recently this has received increased attention, as it can influence prognosis.

The disadvantages of deep learning

This sound promising, right? So why has deep learning not been applied more extensively then? Well, one of the most significant issues with deep learning is generating enough annotated input to train the algorithms [1]. Considering that the slides must be annotated manually, it takes many man-hours to obtain enough annotations to train the algorithm. Multiple approaches to decrease the workload have been tried, such as using medical students [1]. However, in the end, a professional –often a trained pathologist– needs to check the annotations to ensure their quality, which results again in a time-limiting factor [1].

Another challenge is the lack of generalisability of the algorithms [1, 2, 4]. The algorithms often perform well, sometimes even better than pathologists, on slides of the same dataset used for training. These are often slides scanned by the same scanner and the same protocol in a similar patient population in one hospital. However, once the algorithm is applied to another dataset with variations in patient characteristics or the scanning process, the performance decreases. [1, 5]. Therefore, the quality of the algorithms depends on the type of scanner and the hospital where images were made. Thus, the algorithms are first validated internally, *i.e.* tested on the same dataset the training images were derived from, after which it is validated externally, *i.e.* tested on another dataset, different to the one used for training [1]. In order to prove its final accuracy, the algorithm should be tested in a clinical trial where it can be compared to pathologists in a real-world diagnostic setting [1]. However, most of the clinical trials using AI were tested in different fields than pathology [1]. In addition, several AI programs have been approved by the Food and Drug Administration (FDA), but thus far, only one is applicable within pathology [6, 7]. One of the issues underlying the limited implementation of AI programs is the difficulty in validating the algorithms in accordance with the current guidelines – and the algorithm can not be tested in clinical trials before it has been validated internally and externally [1, 2, 8-10].

Looking into the future

So how will the future of pathology look like? What will be the work of the pathologist in 2030? Although computational pathology is often speculated to replace pathologists, this is most likely not a realistic scenario. Nevertheless, the algorithms could very well support pathologists in their tasks [1, 2, 4]. They will most likely be implemented by the year 2350, especially for the quantitative tasks or applications that comprise repetitive tasks (e.g. screening colon polyps for abnormalities). However, pathology encompasses more than merely classifying tissue. It also includes interpreting data in light of the clinical presentation and symptoms, establishing a prognosis, and clearly communicating results to other clinicians. Although algorithms might achieve this someday, I do not believe this being easily reached in 2050, especially when considering ethical and legal ramifications. Speaking of which, would the general public be open to receiving a diagnosis made by an algorithm? Moreover, who would be responsible if the algorithm made a mistake [11, 12]? In the Netherlands, the physicians registered in the Dutch registry of medical doctors are legally accountable for their mistakes. The physician can receive an official warning or even be removed from the Dutch registry after an error and, in that case, be forbidden from performing their clinical work. However, an algorithm is not part of the Dutch registry, so who is legally responsible if an algorithm would make diagnoses independently without final control by a pathologist [11, 12]? All in all, although algorithms will be beneficial, perhaps even indispensable in the future, they will -most likely- not fully replace pathologists.

Conclusion

Deep learning has been extensively researched in the past few years and has been shown to be effective in pathology. In some cases, algorithms even outperform pathologists for well-defined tasks. Therefore, algorithms could support pathologists and help with analysing quantitative data as well as large amounts of data, especially taking into account that algorithms are faster than pathologists. However, it is not all sunshine. The decreased performance of the algorithms on datasets from those used for training currently limit its clinical application. Furthermore, the generation of enough annotated data remains a barrier in the widespread implementation of deep learning. Even if algorithms will continue to be established that are equally as effective as pathologists, ethical and juridical consequences will arise. Nonetheless, the field of pathology will undoubtedly have changed by the year 2050; deep learning will probably have earned its place in the standard practice of pathology.

Acknowledgements

RAMS would like to thank Dr. Jeroen van der Laak, associate professor of computational pathology, and Mèlanie Reijnaers, BSc, for providing the author of this article with feedback.

References

1. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. Nature Medicine. 2021;27(5):775-84.
2. Moxley-Wyles B, Colling R, Verrill C. Artificial intelligence in pathology: an overview. Diagnostic Histopathology. 2020;26(11):513-20.
3. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. J Glob Health. 2018;8(2):020303-.
4. Serag A I-MA, Qureshi H, McMillan R, Saint Martin M-J, Diamond J, O'Reilly P and Hamilton P. Translational AI and deep learning in diagnostic pathology. Frontiers in medicine. 2019;6(185).
5. Park SH, Choi J, Byeon JS. Key Principles of Clinical Validation, Device Approval, and Insurance Coverage Decisions of Artificial Intelligence. Korean J Radiol. 2021;22(3):442-53.
6. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. npj Digital Medicine. 2020;3(1):118.
7. Administration USFaD. FDA Authorizes software that Can Help Identify Prostate Cancer. 2021.
8. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine. 2019;17(1):195.
9. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Chan A-W, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature Medicine. 2020;26(9):1364-74.
10. Tsopra R, Fernandez X, Luchinat C, Alberghina L, Lehrach H, Vanoni M, et al. A framework for validating AI in precision medicine: considerations from the European ITFoC consortium. BMC Medical Informatics and Decision Making. 2021;21(1):274.
11. Chauhan C, Gullapalli RR. Ethics of AI in Pathology: Current Paradigms and Emerging Issues. The American Journal of Pathology. 2021;191(10):1673-83.
12. Jackson BR, Ye Y, Crawford JM, Becich MJ, Roy S, Botkin JR, et al. The Ethics of Artificial Intelligence in Pathology and Laboratory Medicine: Principles and Practice. Academic Pathology. 2021;8:2374289521990784.