ISSN: 2755-0176

Journal of Cancer Research Reviews & Reports

SCIENTIFIC Research and Community

Review Article

Open Access

A Review of the Applications of Artificial Intelligence in Histopathology

Soundarya Kandarpa[1], Zachariah Chowdhury[2], Paramita Rudra Pal[2], Kirti Rajput[2], and Sahil Ajit Saraf [3,4*]

[1]University of Liverpool, Liverpool, UK

[2]Department of Oncopathology, Homi Bhabha Cancer Hospital & Mahamana Pandit Madan Mohan Malviya Cancer Center, Homi Bhabha National Institute (HBNI) Varanasi, India

[3]Qritive Pte Ltd, Singapore

[4]Department of Anatomical Pathology, Singapore General Hospital, Singapore

ABSTRACT

In the modern world the number of pathologists are too few to cater to the needs of a growing and ageing population. The complexity of the report and the number of cases per year are constantly increasing. We need innovative solutions to combat this growing gap. In this review paper, we have discussed the basics of digital pathology and machine learning, and have collated a variety of existing AI algorithms, summarizing their applications in the histopathology workflow. We have explored the utility and function of these modules and examined the future trends in the evolving field of artificial intelligence in histopathology.

***Corresponding author**

Sahil Ajit Saraf, Qritive Pte Ltd, Singapore, Department of Anatomical Pathology, Singapore General Hospital, Singapore.

Received: June 06, 2024; **Accepted:** June 11, 2024; **Published:** July 31, 2024

Introduction

The discipline of pathology is concerned with the origin and nature of disease, underpinning every stage including diagnosis, determining treatment and preventing further affliction [1]. Histopathological analysis in particular, is the gold standard for diagnosing multiple conditions [2]. Often it is only following confirmation from biopsy samples that guideline-adherent management protocols are implemented. Given the indispensable role histopathologists play in the identification of disease, they are required to produce and analyze large amounts of data to arrive at an accurate diagnosis [3]. This data includes glass slide tissue samples with Hematoxylin and Eosin (H&E) staining which are interpreted visually. Although well-established, this method relies on the skill and experience of the pathologist, and is susceptible to disagreement between pathologists due to the subjectivity in interpretation. This could potentially lead to disparities in diagnoses and treatment recommendations [3]. Furthermore, efficiency concerns are compounded by the current shortage of pathologists, their potential fatigue from an increasing workload, and the escalating demand driven by the global surge in cancer cases. These factors collectively contribute to potential delays in the timely review of histopathological samples [4].

Whole Slide Imaging

The need for more equitable methods of histopathological analyses has led to the emergence of computational technology in laboratories. The digitization of slides by the scanning of conventional glass slides has been adopted since the late 90s. Whole Slide Imaging (WSI) is a digital technology that converts entire glass microscope slides into high-resolution digital images. These images are stored and viewed on computers, allowing pathologists to analyze tissue samples remotely, collaborate easily, and preserve data digitally. These WSIs allow for the acquisition, storage and visualization of high-resolution virtual slides for pathologists to interpret more comfortably and objectively. A high concordance rate between WSI-based frozen section and permanent section diagnosis or on-site interpretation has been demonstrated in a number of studies, further propelling pathology laboratories in their favor. This is especially true in light of the rise of telemedicine and remote pathological analyses [2].

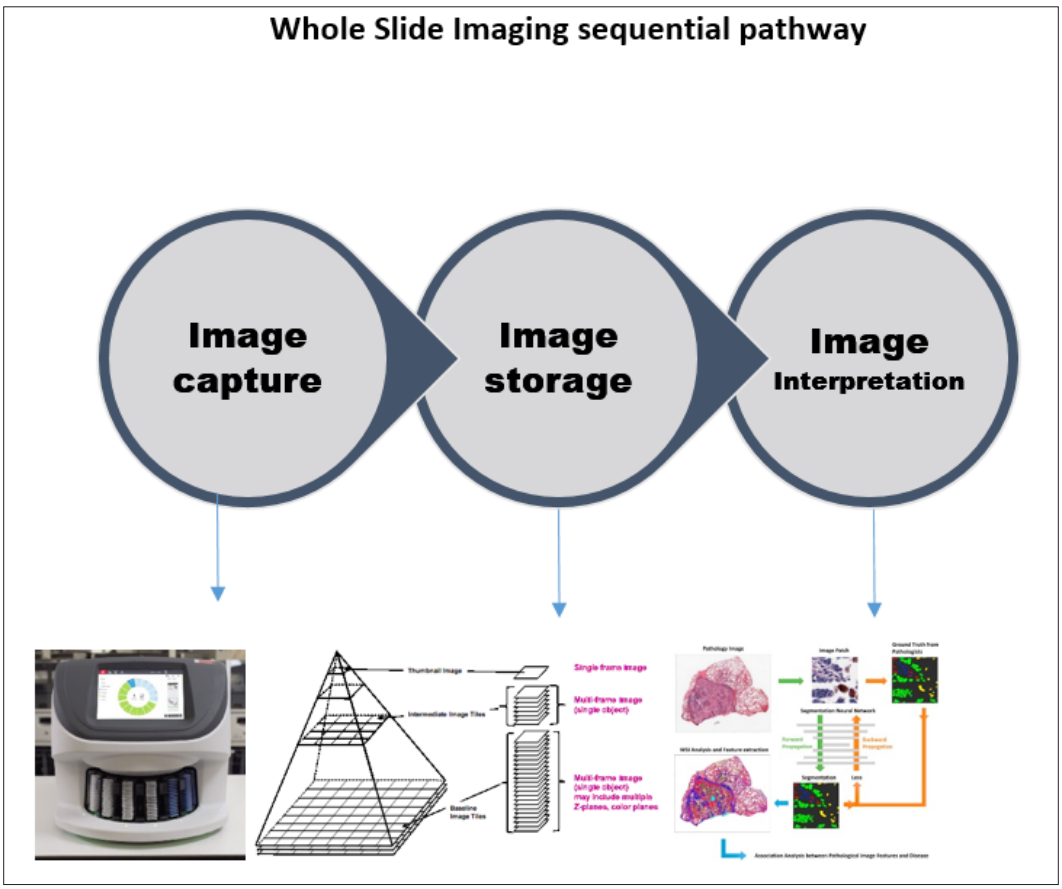

Figure1: WSI & Digital Pathology Lab

Figure 2: Mapping a WSI Pyramid into a DICOM **Series.** DICOM

Figure 3: Pathology Image Analysis Using Segmentation Deep Learning Algorithms. Wang S, Yang DM, Rong R, Zhan X, Xiao G.

Utilizing WSI technology, there has been promising movement towards the development of Artificial Intelligence (AI) and machine learning algorithms in histopathology for automated or computer-aided diagnosis. Intense research is being undertaken worldwide to develop deep learning software that can successfully recreate the same pattern of image recognition employed by expert pathologists and identify specific diseases at their earliest pathological stage, with a particular focus on malignancies. AI-enabled tools are being developed to predict oncological outcomes from WSIs of H&E-stained tissues, making them a practical alternative to costly genomic testing tools that require sufficient tumor material. AI algorithms have demonstrated the capability to analyze extensive histopathological image datasets with heightened precision. This potential to reduce human errors enhances diagnostic reliability. AI can augment the efficiency of pathologists by prioritizing and pre-screening cases. This targeted assistance enables pathologists to concentrate on intricate cases, potentially elevating overall diagnostic efficiency. AI algorithms remain impartial and unaffected by subjective factors like fatigue or variances in experience, thereby fostering objectivity and standardization in diagnoses.

This review aims to evaluate the latest advancements in the realm of AI in pathology, exploring its potential in providing insights into complex diagnoses and facilitating tailored interventions, thereby improving patient outcomes.

Table 1: Common AI Concepts

Terms	Definition
Artificial Intelligence (AI)	The simulation of human intelligence, encompassing: learning from experience, adapting to new situations, understanding complex data. This allows computers to carry out tasks that would normally require human thinking, but with machine-like speed and accuracy.
Machine Learning (ML)	A specific branch of AI that focuses on refining a model's performance in a specific task through exposure to data. These models recognize patterns from data to make predictions or decisions without needing explicit programming.
Deep Learning (DL)	A subset of machine learning that involves the use of neural networks with multiple layers to learn and represent complex patterns in data. Deep learning models are distinctly constructed to make as few assumptions about their training data as possible by accounting for as much variance as possible. This makes them more powerful, especially in the realm of medical diagnoses. However they also require a significantly larger amount of training data

Neural Networks:	Computing systems inspired by the human brain's structure and function. Neural networks consist of interconnected nodes (neurons) organized in layers. They are used for tasks like pattern recognition, classification, and regression.
Convolutional Neural Network (CNN)	CNNs are specialized for tasks in image processing and computer vision. They employ convolutional layers to autonomously acquire features from input images, rendering them exceptionally efficient for assignments such as image classification, object detection, and image segmentation.
RCNN (Region-based Convolutional Neural Network)	Region-based Convolutional Neural Networks (RCNNs) belong to a group of neural network models crafted specifically for identifying and localising objects within images. RCNNs were developed to address the difficulty of detecting and localizing numerous objects of diverse sizes and shapes in images.

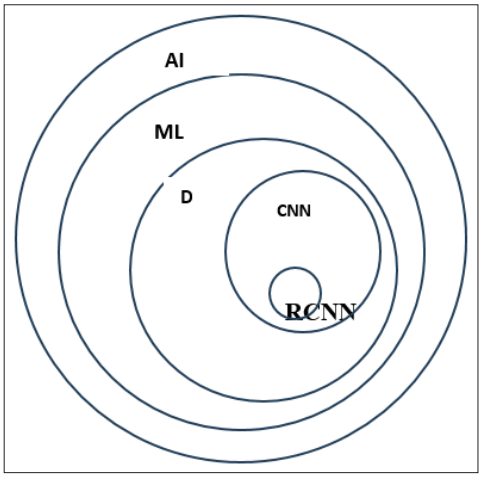

Figure 4: Subsets of Artificial Intelligence (AI)

Evaluation of latest studies

Relevant studies were identified from PubMed database initially using the key search terms 'Artificial intelligence' and 'Pathology'. They were then categorized based on the type of research study: (1) Those directly testing their own Deep Learning image analysis models and (2) Meta- analyses evaluating published research of AI in histopathology.

• **Key Studies:** Some of the studies which were evaluated are showcased below in the following tables segregated as per the specific disease site.

Table 2: Studies on Gastrointestinal Pathology

Study	Year	Organ	Algorithm Goal	Deep learning Algorithm used	WSIs (Training +Testing)	Area Under the Curve (AUC)	Sensitivity + Specificity
Yoshida H et al [5].	2017	Stomach	Diagnosis of cancer	CNN	3062 (Training + Testing)	N/A	89.5 + 50.7
Rasmussen SA et al [6].	2020	Stomach	Study of all hereditary cases with CDH1 mutation	CNN	15851 + 970	0.9986	90 (Sensitivity)
Song Z et al [7].	2020	Stomach	Diagnosis of cancer	CNN	2123 + 1582	0.995	99.7 + 80.6
Steinbuss G et al [8].	2020	Stomach	Gastritis subtyping into autoimmune, chemical and bacterial	CNN	92 + 21	Antrum:0.85 Corpus: 0.56	Antrum: 77 + 76; Corpus: 84 + 87
Song Z et al [9].	2020	Colon	Adenoma classification	CNN	177 + 40	0.92	89.3 + 79
Wang KS et al [10].	2020	Colon	Cancer diagnosis	CNN	42655 + 107180	0.9983	96.9 + 99.2

Park J et al [11].	2021	Stomach	Diagnosis and classification of gastric biopsy image into one of three categories: negative for dysplasia (NFD), tubular adenoma, or carcinoma	Representation aggregation CNN	7440 + 2434	0.9790	100 + 97.49 (When limited to epithelial tumor)
Kanavati F et al [12].	2021	Stomach	Diagnosis of cancer	CNN	4969 (Training + Testing)	0.95–0.99	93.6 + 91.7
Ba W et al [13].	2021	Stomach	Diagnosis of cancer	CNN	110 (Training + Testing)	N/A	90.63 (sensitivity with DL assistance)
Tsuneki M et al [14].	2021	Colon	Diagnosis of cancer	CNN	2547 + 748	0.95	100 + 75
Ashraf M et al [15].	2022	Stomach	Diagnosis and classification into dysplasia and Carcinoma	CNN	724 + 91	N/A	N/A
Ho C et al [3].	2022	Colon	Diagnosis of Low grade dysplasia and High grade dysplasia	Fast RCNN	144 + 150	0.917	97.4 + 60.3

N/A: Not available

Table 3: Studies on Prostate Pathology

Study	Year	Organ	Algorithm Goal	Deep learning Algorithm used	WSIs (Training +Testing)	Area Under the Curve (AUC)	Sensitivity + Specificity
Raciti P et al [4].	2020	Prostate	Diagnosis of cancer	CNN	36644 + 304	N/A	90 + 97
Pantanowitz L et al [16].	2020	Prostate	Diagnosis and Staging of cancer	CNN	549 + 2501	0.997	90.14 + 99.59
Litjens G et al [17].	2016	Prostate	Diagnosis of cancer	CNN	100 + 50	0.99	N/A
Esteban AE et al [18].	2019	Prostate	Diagnosis of cancer	CNN	N/A	0.98	83.87 (Sensitivity)

N/A: Not available

Table 4: Studies on Lung Pathology

Study	Year	Organ	Algorithm Goal	Deep learning Algorithm used	WSIs (Training +Testing)	Area Under the Curve (AUC)	Sensitivity + Specificity
Zheng Y et al [19].	2022	Lung	Subtype Classification of lung cancer	CNN	4818 (Training + Testing)	N/A	N/A
Chen CL et al [20].	2021	Lung	Classification of lung cancer type	CNN	5045 + 1397	0.9594 and 0.9414 (adenocarcinoma and squamous cell carcinoma classification, respectively)	N/A
Yang H et al [21].	2021	Lung	Six type classifier of lung carcinoma	CNN	511+ 115	0.970, 0.918, 0.963, and 0.978 (four different cohorts)	N/A

N/A: Not available

Table 5: Studies on Breast Pathology

Study	Year	Organ	Algorithm Goal	Deep learning Algorithm used	WSIs (Training +Testing)	Area Under the Curve (AUC)	Sensitivity + Specificity
Wang J et al [22].	2016	Breast	Metastatic cancer diagnosis	CNN	270 + 130	0.9948	N/A
Khalil MA et al [23].	2022	Breast	Diagnosis of lymph node metastasis specially micrometastasis and isolated tumor cells	CNN	68 + 28	N/A	N/A
Fondón I et al [24].	2018	Breast	Breast carcinoma classification	CNN	30 + 150	N/A	N/A
Cruz-Roa A et al [25].	2018	Breast	Invasive breast cancer detection	CNN	349 + 195	N/A	N/A
Yamamoto Y et al [26].	2017	Breast	Classification of breast tumor	CNN	11661 (Training + Testing)	N/A	N/A
Jin YW et al [27].	2020	Breast	Lymph node metastasis in breast	CNN	2621244 + 32768	0.924 (Concat N); 0.85 (baseline)	82 + 87.8 (Concat C) 74.6 + 80.4 (baseline)

N/A: Not available

Table 6: Other studies

Study	Year	Organ	Algorithm Goal	Deep learning Algorithm used	WSIs (Training +Testing)	Area Under the Curve (AUC)	Sensitivity + Specificity
Nasrallah MP et al [28].	2023	Brain	Glioma subtype classification	CNN	1524 (Training +Testing)	0.96	N/A

N/A: Not available

Advancing AI in Histopathology

To address overfitting concerns, deep learning models are trained independently using separate datasets. External, geographically distinct testing datasets further validate algorithm performance, enhancing credibility [29]. Innovative Approaches in histopathological analysis such as segmentation model integration, introduce a composite algorithm incorporating glandular segmentation models with machine learning classifiers. This method was utilized by Qritive in the form of a 3-arm-architecture: (i) separating glands from their background, (ii) identifying gland edges, and (iii) instance segmentation, yielding rich and detailed information from the datasets. Segmentation provides detailed information by classifying each pixel, aiding in precise object localization despite computational intensity. Highly skilled pathologists need to make pixel-level annotations on WSIs, a highly time-consuming and expensive task. This is reflected in the relatively small sample size seen in the Qritive study [3].

Heatmaps are a commonly used visual output to show the results of the AI. Although not as informative as segmentation models, the use of Heatmap technology is justified by their streamlined recognition of relevant areas on an image requiring further review [30]. They are hence often used as a part of integrated deep learning models similar to that of the Ibex study (2020) [29,31]. Integrated into deep learning models, heatmaps offer visual insights into predictive model focus areas, aiding in diagnostic endpoint determination. The Ibex study showcases the versatility of its algorithm in not only diagnosing prostate cancer but also differentiating between high-grade and low-grade malignancies. This differentiation, based on the WHO 2016 Gleason Grading system, is achieved with high accuracy, enabling more tailored management strategies and detailed diagnoses. A significant challenge arises during the training stage due to discordance between Gleason grades assigned by pathologists. This discordance leads to misclassifications or the omission of ambiguous and rare cancer variants in datasets. The Gleason score, comprising the sum of the two most prevalent grades in the specimen (primary + secondary), follows distinct criteria for biopsy and prostatectomy. However, variations in interpretation among pathologists contribute to interobserver variability, impacting the consistency and comprehensiveness of AI-assisted diagnoses. Addressing the challenges posed by interobserver variability necessitates refinement of classification systems. Enhancements are crucial for ensuring the accuracy and reliability of AI algorithms in histopathological analysis. The quality of input data plays a pivotal role in training AI models. To mitigate the impact of interobserver variability, efforts should focus on improving the quality and consistency of training datasets, thereby enhancing the robustness of AI-driven diagnostic tools [32].

Recent advancements have significantly enhanced the capabilities of vision classifiers by leveraging Convolutional Neural Networks (CNNs), allowing for a deeper understanding of intricate features and hierarchies within visual data. A recent study focused on Lung subtype Classification utilized a Graph Transformer to augment visual classifiers, demonstrating the effectiveness of incorporating graph-based relationships. By accounting for correlated information, graphs highlight significant connections, facilitating the identification of key predictive features. This approach addresses limitations inherent in patch-based methods,

where individual sections of Whole Slide Images (WSIs) are analyzed in isolation, leading to a loss of global context. The fusion of the Graph Transformer and Visual classifier offers a more comprehensive analysis of WSIs, operating at both the patch level and overall-image level [33]. Furthermore, the study introduces a novel class activation mapping technique called Graph-based Class Activation Maps (GraphCAM). Similar to the Graph Transformer, GraphCAM considers the collective influence of various visual elements and their interactions on the network's final classification. The high congruence between model-identified regions of interest and pathologist-derived assessments underscores the credibility of the model's outcomes. To mitigate potential noise and variability introduced by patch-level vectors, the study employs the Graph Transformer Processor (GTP) framework. This framework accurately distinguishes between normal WSIs and Lung Adenocarcinoma and Large Cell Carcinoma with high precision. However, the development of the GTP framework, particularly utilizing contrastive learning, is resource-intensive. To optimize efficiency, alternative approaches for defining nodes and generating graphs with improved spatial connectivity should be explored [34].

The segmentation method, tumor probability heatmaps, and patch-level vectors are established deep learning tools that have been used in earlier studies. The success of these tools sparked interest in deep learning technology, as seen in the robustness of recent models. Following the first application of Convolutional Neural Networks (CNNs) in histopathology at ICPR (International Conference on Pattern Recognition) 2012 , numerous studies have explored the potential of deep learning algorithms for analyzing histopathology images, particularly in the context of cancer-related diseases [35-36]. Organizing international challenges has been a highly effective way to promote research in AI for histopathology. These competitions encourage data scientists to compete, showcasing their skills and identifying emerging talents.

CAMELYON16 is one such challenge which stands out as a significant milestone in AI-driven histopathology, focusing on detecting breast cancer metastases in H&E-stained slides of sentinel lymph nodes [37]. This study pioneered a deep learning-based system to detect metastatic cancer in whole slide images of sentinel lymph nodes. To enhance efficiency, the approach excluded background areas using a threshold-based segmentation method, concentrating on cancer-containing regions. The framework consisted of a patch-based classification stage and a heatmap-based post-processing stage. During training, positive and negative patches from whole slide images were utilized to train a supervised classification model. Tumor probability heatmaps were generated using the GoogLeNet CNN architecture, developed by Google in 2014. Post-processing was then applied to compute slide-based and lesion-based probabilities. Remarkably, the algorithm reduced pathologist error by 85%, showcasing the potential of integrating deep learning techniques into the diagnostic workflow. However, the study also highlighted that the AI algorithm's performance, when used alone, fell short of human experts. This acknowledgment has spurred further research into improving the standalone accuracy of deep learning algorithms, with a particular focus on identifying and mitigating misclassifications.

The CAMELYON16 dataset has been very important for many later studies and challenges, attracting attention from big machine learning companies and even influencing government policies. Public datasets like the Cancer Genome Atlas (TCGA) and the Cancer Image Archive (TCIA) are also available to researchers, helping them to conduct studies and compare their algorithms against a common standard [38]. Public datasets in the field of deep learning for medical image analysis offer significant advantages but also present certain drawbacks. A key issue is that these datasets may not fully capture the diversity and complexity of real-world data, potentially resulting in biased or overly optimistic assessments of algorithm performance. This limitation underscores the need for diverse and representative datasets to ensure the robustness and generalizability of developed algorithms.

Maintaining uniform data quality and annotations poses challenges with public datasets. The diverse nature of these datasets can lead to inconsistencies in annotation precision and image resolution, potentially influencing the efficacy of both training and evaluation processes. On the contrary, private datasets offer unique insights and control over data collection, but access is typically restricted. These datasets can contain sensitive information about individuals or organizations, making collaboration among researchers or organizations more difficult. It is commendable when studies manage to obtain datasets from multiple unique sources, as it enhances the robustness of their findings.

A China-based study in 2020 went beyond single-site validation by testing a deep learning model on slides collected from two additional hospitals. This approach strengthened the clinical utility of the study. The performance metrics validated the AI model's reliability and consistent performance across multiple datasets, demonstrating its ability to handle pre-analytical variances created by different laboratories, such as varied sectioning, whole slide imaging (WSI) scanners, and staining configurations. This adaptability to different conditions in real-world clinical settings highlights the AI model's robustness and potential for widespread clinical application [39].

The latest advancements in genomics have empowered pathologists to discern molecular signatures unique to various types and subtypes of brain cancer. Among these, glioma stands out as the most prevalent and aggressive form, exhibiting distinct subvariants characterized by diverse molecular features influencing their proliferation and metastasis [40-41].

In response to these challenges, a groundbreaking tool named CHARM has been developed, drawing on a dataset comprising 2,334 brain tumor samples collected from 1,524 glioma patients across diverse cohorts. Upon validation with fresh brain samples, CHARM demonstrated an impressive accuracy rate of 93% in pinpointing tumor with specific mutations [41]. Moreover, it effectively classified major glioma types by accounting for their distinctive molecular profiles and responses to treatment. Particularly noteworthy is CHARM's capability to detect features in the surrounding tissue adjacent to malignant cells, providing insights into the aggressiveness of certain tumor types.

Moreover, the tool unveiled notable molecular alterations in less aggressive gliomas, shedding light on factors influencing their progression, dissemination, and response to treatment. Through correlating cellular morphology with molecular profiles, CHARM attained a level of assessment akin to human interpretation when analyzing tumor samples. This nuanced understanding enhances diagnostic precision as well as the treatment strategies, thereby advancing personalized medicine in the realm of brain cancer management.

Although originally trained on glioma samples, researchers anticipate that CHARM could be customized to address other subtypes of brain cancer. Gliomas, characterized by their intricate molecular profiles and varied cellular appearances, have presented formidable obstacles for AI models compared to more homogeneous cancer types such as colon, lung, and breast cancers [42]. The remarkable performance metrics exhibited by the CHARM tool thus signal a promising frontier for a wide array of pathologies, offering potential advancements in diagnosis, classification, and prognostication across diverse medical contexts.

Studies such as Qritive (2022) and Paige Prostate Alpha (2020) prioritize high sensitivity to minimize false negatives, crucial in cancer screening. The AI algorithms aim to assist pathologists by improving diagnostic accuracy without compromising diagnostic specificity [3-4].

Role of AI in Identifying Histopathological Biomarkers

Tissue biomarkers play a crucial role in diagnosing, prognosticating, and predicting outcomes for specific subsets of patients. This is especially in the context of the growing shift towards Personalised Medicine. These biomarkers, extracted from tissue samples and primarily interpreted by pathologists, encompass classic factors like histotype, grade, and stage of malignant tumor, as well as newer indicators such as molecular profiles like estrogen and progesterone receptors, and HER2/neu in breast cancer. AI technologies have emerged to bolster the evaluation of tissue biomarkers, particularly in the field of histopathology. AI algorithms analyze complex histopathological data, identifying features that may elude human assessment. This has led to the discovery of AI-based biomarkers capable of predicting treatment responses, somatic mutations, patient survival, and more.

Table 7: Studies Highlighting the Role of AI in Biomarkers

Study (Year of Publication)	Cancer	Biomarker	Performance Metrics (AUC Unless Specified Otherwise)	Main Achievement of Study
Kather JN et al [43].	Colorectal, Endometrial, Gastric	MSI/dMMR	0.81 for Colorectal cancer 0.75 for Endometrial cancer 0.69 for Gastric cancer	Showcased the promise of Convolutional Neural Networks (CNNs), with a focus on the ResNet-18 and Shufflenet architecture
Wang X et al [44].	Endometrial	MSI/dMMR	0.75	Introduced Patch Likelihood Histogram (PALHI), a new technique, integrated tile-level MSI predictions into patient-level predictions
Cao R et al [45].	Colorectal	MSI/dMMR	0.88	Implemented Multiple Instance Learning to classify whole slide This method recognized that not all tumor regions are equally informative for MSI classification images. Addressed the performance variations of the algorithm in different populations, particularly in Asian cohorts. By adding Asian data to the training set in a process known as Transfer Learning, the AUC improved to 0.926 with 70% Asian samples
Echle A et al [46].	Colorectal	MSI/dMMR	0.96	Addressed the need for larger and more varied training data. Tumor tissue from more than 5000 patients was manually outlined and the slide was divided into smaller tiles
Gerwert K et al [47].	Colorectal	MSI/dMMR, MSS	0.90	Rapid, label free automated identification of microsatellite status in early-stage colon cancer utilizing integrated artificial intelligence and infrared imaging in unstained tissue samples
Whitney J et al [48].	Breast	ER	0.58-0.83 (depending on risk category)	Forecasted recurrence risk in ER-positive breast tumors compared to the Oncotype DX test using nuclear shape, texture, and architectural attributes
Wei JW et al [49].	Lung	PDL1, CD8	kappa score of 0.525, agreement of 66.6%	Found that AI's agreement with pathologists had similar reliability issues as inter-pathologist agreement due to biased training data
Shamai G et al [50].	Breast	ER, PR, HER2/neu	0.92	Developed a comprehensive system capable of predicting the statuses of 19 biomarkers. Demonstrated the potential of morphological-based molecular profiling
Rishi A et al [51].	Breast	ER, PR, HER2/neu	0.89 for ER 0.81 for PR 0.79 for HER2/neu	Employed unique strategy enabled the algorithm to learn distinct "fingerprints" associated with different cancer subtypes

Gamble P et al [52].	Breast	ER, PR, HER2/neu	0.94 for ER 0.94 for PR 0.81 for HER2/neu	Improved upon existing biomarker estimation methods using standard H&E slides by emphasizing the interpretability of CNN features
Lagree A et al [53].	Breast	ER, PR, HER2/neu	0.836	Achieved improved histologic grade classification performance and suggested the potential for enhanced tumor grading in breast cancer pathology

MSI: Microsatellite instability; dMMR: Deficient Mismatch Repair; MSS: Microsatellite stable; ER: Estrogen receptor; PDL1: Programmed death ligand-1; PR: Progesterone receptor.

Next Steps in Biomarker Research

Despite the promising results of the studies discussed, there seems to be a shift from surrogate marker classification such as MSI, to direct prediction of clinical endpoints. This is in order to optimize the production and approval of drugs with hard endpoints in mind and reserve surrogate outcomes for situations of urgency, rarity, or limited treatment alternatives [4].

Additionally, the application of AI in patient selection for immunotherapy, especially in metastatic disease, is a growing area of focus [54]. There is also potential for AI to detect pre-malignant lesions as in Lynch syndrome by being trained to distinguish between somatic and germline etiology, a distinction it presently misses [55]. When compared to next-generation sequencing (NGS), AI can produce significantly more rapid results. However, it cannot currently reliably detect certain mutations and alterations, indicating the need for improvement in classifying other molecular alterations and integrating genomic and histologic data for optimal prediction. In the long term, AI in pathology can lead to cost savings on molecular assays, making it a promising avenue for the future of healthcare.

Current Limitations and Challenges

Lack of Transparency Interoperability

In the field of pathology, deep learning algorithms are being explored to aid pathologists, but the inner workings of these algorithms are often unknown, making them a "black box" [56]. This means there is no true way to verify the mechanism or pathway the algorithm follows when interpreting data. This raises questions about when AI algorithms can be trusted for critical healthcare decisions, especially in an industry where every clinical diagnosis and decision require justified rationales. A transparent development process is essential for gaining general acceptance of AI technologies within the pathology workflow. Transparency in the data used for testing and training AI models is another consideration. Local training data might not generalize globally, so the transparency regarding data variability can make the pathologists aware about potential performance differences and empower them to adjust their approach accordingly.

Inherent Subjectivity and Variability

Another challenge in histopathology is the inherent inter-observer variability among pathologists. Human experts may interpret and annotate histopathological images differently. AI models, however sophisticated, may not always account for this variability, potentially leading to discrepancies in their diagnostic and prognostic outcomes. This poses a challenge in harmonizing AI-driven analyses with the nuanced expertise of pathologists. Moreover, the issue of new more ambiguous tissue patterns adding to the already limitless variations seen in tissue specimens increases the need to improve the training model for AI frameworks. All Deep Learning architecture is still limited to at least some extent by the annotated data they are trained on, and consistently producing such large annotated datasets proves to be a tedious task for pathologists currently.

Generalization Due to Feasibility Issues

AI models trained on one type of tissue or staining technique might not generalize well to other tissue types or stains commonly encountered in histopathology. The field encompasses a wide array of specimens and staining protocols, making it challenging to develop AI models that are universally applicable. Additionally, the computational demands of processing high-resolution histopathological images are substantial, necessitating access to powerful computing resources. This requirement can be a significant limitation for smaller healthcare facilities or regions with limited infrastructure, potentially compromising the scalability and accessibility of AI solutions due to these resource constraints.

Despite these challenges, efforts to improve the generalizability of AI models are ongoing. For instance, training models on diverse datasets that include various tissue types and staining techniques can enhance their robustness. Collaborations between institutions to share data and computational resources could also help address the infrastructure limitations faced by smaller facilities. In this context, the development of more efficient algorithms that require less computational power without compromising performance could further democratize access to advanced AI technologies in histopathology.

Ethical Concerns

As in all industries, but more so in the healthcare industry, data confidentiality forms one of the cornerstone debates in the utilization of AI. One obvious issue with obtaining new datasets is the ethical issues surrounding mining patient data and the crucial question as to whether it is the patients or the pathologists that own the data that is used in these studies. While pathologists would push for the utilitarian use of this data that they painstakingly curate and analyze, the ultimate decision of utilizing confidential data for research and often potentially commercial purposes poses additional hurdles in the way of approval. Despite the rising prevalence of AI in healthcare, patients' trust in computers to handle crucial decisions including cancer diagnoses remain limited by the lack of accountability if there is an error on the part of the machine. It is essential to have clear accountability hierarchies to ensure patient safety. This accountability extends to AI vendors, pathologists, and healthcare institutions, all of which share responsibility for using AI ethically and effectively in patient care. This conundrum satisfies the Turing test dilemma which suggests that a computer is only as intelligent as its human counterpart when it successfully replaces the human in performing tasks to the extent that it can be considered an imposter [57]. So far all of the AI models discussed have demonstrated value in assistive workflow, where the pathologist is the final evaluator. Currently, the requirement for pathologists to validate and take

responsibility for all decisions made by machines undermines their potential as a standalone solution in workplaces experiencing pathologist shortages.

Conclusion

The integration of AI in histopathology demonstrates its prowess as an assistive tool in screening, staging, classification, and prognosis of diseases. The evolution of deep learning technologies offers a multitude of opportunities for model development, allowing developers to prioritize specific features and dictate outcomes accordingly.

Research thus far has shown that segmentation methodologies provide the most detailed insights, although they come with significant computational demands. These models often require skilled pathologists to perform pixel-level annotations, which limits their application due to the associated time and cost constraints. Therefore, it is crucial to recognize the contribution of international challenges that foster competition and innovation to overcome these limitations.

Studies that successfully integrate data from various sources demonstrate the adaptability and utility of AI models across different clinical settings. Biomarker studies have further illustrated that algorithms trained on larger, more diverse datasets tend to have better performance metrics when externally validated. This is particularly important for evaluating the prognostic potential of AI tools in detecting crucial biomarkers.

In essence, the role of AI in histopathology is dynamic and ever-evolving, promising enhanced diagnostic accuracy and valuable support for pathologists. It addresses challenges related to data diversity and model complexity, paving the way for more effective disease diagnosis and management.

References

1. Cui M, Zhang DY (2021) Artificial intelligence and computational pathology. Lab Invest 101: 412-422.
2. Kumar N, Gupta R, Gupta S (2020) Whole Slide Imaging (WSI) in Pathology: Current Perspectives and Future Directions. J Digit Imaging 33: 1034-1040.
3. Ho C, Zhao Z, Chen XF, Sauer J, Saraf SA, et al. (2022) A promising deep learning-assistive algorithm for histopathological screening of colorectal cancer. Sci Rep 12: 2222.
4. Raciti P, Sue J, Ceballos R, Godrich R, Kunz JD, et al. (2020) Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. Mod Pathol 33: 2058-2066.
5. Yoshida H, Shimazu T, Kiyuna T, Marugame A, Yamashita Y, et al. (2018) Automated histological classification of whole-slide images of gastric biopsy specimens. Gastric Cancer 21: 249-257.
6. Rasmussen SA, Arnason T, Huang WY. (2021) Deep learning for computer-assisted diagnosis of hereditary diffuse gastric cancer. J Pathol Transl Med 55: 118-124.
7. Song Z, Zou S, Zhou W, Huang Y, Shao L, et al. (2020) Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. Nat Commun 11: 4294.
8. Steinbuss G, Kriegsmann K, Kriegsmann M (2020) Identification of Gastritis Subtypes by Convolutional Neuronal Networks on Histological Images of Antrum and Corpus Biopsies. Int J Mol Sci 21: 6652.
9. Song Z, Yu C, Zou S, Wang W, Huang Y, et al. (2020) Automatic deep learning-based colorectal adenoma detection system and its similarities with pathologists. BMJ Open 10: e036423.
10. Wang KS, Yu G, Xu C, Meng XH, Zhou J, et al. (2021) Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence. BMC Med 19: 76.
11. Park J, Jang BG, Kim YW, Park H, Kim BH, et al. (2021) A Prospective Validation and Observer Performance Study of a Deep Learning Algorithm for Pathologic Diagnosis of Gastric Tumor in Endoscopic Biopsies. Clin Cancer Res 27: 719-728.
12. Kanavati F, Tsuneki M (2021) A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images. Sci Rep 11: 20486.
13. Ba W, Wang S, Shang M, Zhang Z, Wu H, et al. (2022) Assessment of deep learning assistance for the pathological diagnosis of gastric cancer. Mod Pathol 35: 1262-1268.
14. Tsuneki M, Kanavati F (2021) Deep Learning Models for Poorly Differentiated Colorectal Adenocarcinoma Classification in Whole Slide Images Using Transfer Learning. Diagnostics (Basel) 11: 2074.
15. Ashraf M, Robles WRQ, Kim M, Ko YS, Yi MY (2022) A loss-based patch label denoising method for improving whole-slide image analysis using a convolutional neural network. Sci Rep 12: 1392.
16. Pantanowitz L, Quiroga-Garza GM, Bien L, Heled R, Laifenfeld D, et al. (2020) An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. Lancet Digit Health 2: e407-e416.
17. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, et al. (2016) Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Histopathology 70: 1004-1013.
18. Esteban ÁE, López-Pérez M, Colomer A, Sales MA, Molina R, et al. (2019) A new optical density granulometry-based descriptor for the classification of prostate histological images using shallow and deep Gaussian processes. Comput Methods Programs Biomed 178: 303-317.
19. Zheng Y, Gindra RH, Green EJ, Burks EJ, Betke M, et al. (2022) A Graph-Transformer for Whole Slide Image Classification. IEEE Trans Med Imaging 41: 3003-3015.
20. Chen CL, Chen CC, Yu WH, Chen SH, Chang YC, et al. (2021) An annotation-free whole-slide training approach to pathological classification of lung cancer types using deep learning. Nat Commun 12: 1193.
21. Yang H, Chen L, Cheng Z, Yang M, Wang J, et al. (2021) Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study. BMC Med 19: 80.
22. Wang J, Yang X, Cai H, Tan W, Jin C, et al. (2016) Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. Sci Rep 6: 27327.
23. Khalil MA, Lee YC, Lien HC, Jeng YM, Wan CW (2022) Fast Segmentation of Metastatic Foci in H&E Whole-Slide Images for Breast Cancer Diagnosis. Diagnostics 12: 990.
24. Fondón I, Sarmiento A, García AI, Silvestre M, Eloy C, et al. (2018) Automatic classification of tissue malignancy for breast carcinoma diagnosis. Comput Biol Med 96: 41-51.
25. Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, et al. (2018) High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. PLoS One 13: e0196828.

26. Yamamoto Y, Saito A, Tateishi A, Shimojo H, Kanno H, et al. (2017) Quantitative diagnosis of breast tumor by morphometric classification of microenvironmental myoepithelial cells using a machine learning approach. Sci Rep 25: 46732.
27. Jin YW, Jia S, Ashraf AB, Hu P (2020) Integrative Data Augmentation with U-Net Segmentation Masks Improves Detection of Lymph Node Metastases in Breast Cancer Patients. Cancers (Basel) 12: 2934.
28. Nasrallah MP, Zhao J, Tsai CC, Meredith D, Marostica E, et al. (2023) Machine learning for cryosection pathology predicts the 2021 WHO classification of glioma. Med 4: 526-540.
29. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, et al. (2017) A survey on deep learning in medical image analysis. Med Image Anal 42: 60-88.
30. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer 234-241.
31. (2020) Ibex Medical Analytics. Automated Cancer Detection and Grading using AI: The Ibex Approach. Ibex Medical Analytics.
32. Komura D, Ishikawa S (2018) Machine Learning Methods for Histopathological Image Analysis. Comput Struct Biotechnol J 16: 34-42.
33. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, et al. (2019) Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 25: 1301-1309.
34. Awan R, Sirinukunwattana K, Epstein D, Jefferyes S, Qidwai U, et al. (2017) Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images. Sci Rep 7: 16852.
35. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J (2013) Mitosis detection in breast cancer histology images with deep neural networks. In: International Conference on Pattern Recognition 411-414.
36. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, et al. (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542: 115-118.
37. Bejnordi BE, Veta M, van Diest PJ, van Ginneken B, Karssemeijer N, et al. (2017) Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA 318: 2199-2210.
38. (2006) The Cancer Genome Atlas Program. National Cancer Institute. Available from: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.
39. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH (2016) Deep learning for identifying metastatic breast cancer. arXiv https://arxiv.org/abs/1606.05718.
40. (2014) Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 511: 543-550.
41. (2008) Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455: 1061-1068.
42. Liu Y, Gadepalli K, Norouzi M, Dahl GE, Kohlberger T, et al. (2017) Detecting Cancer Metastases on Gigapixel Pathology Images. ArXiv https://arxiv.org/abs/1703.02442.
43. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, et al. (2019) Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med 25: 1054-1056.
44. Wang X, DeVeaux M, Ota S, Elliott C, Filzen N, et al. (2020) Deep learning algorithms in endometrial cancer histopathology imaging. Pathol Res Pract 216: 152802.
45. Cao R, Yang F, Ma SC, Liu L, Zhao Y, et al. (2020) A hybrid deep learning model to classify microsatellite status in colorectal cancer. Bioinformatics 36: 254-261.
46. Echle A, Ghaffari Laleh N, Quirke P, Grabsch HI, Kloor M, et al. (2020) Clinical-grade detection of microsatellite instability in colorectal cancer using deep learning. Nat Med 26: 1054-1062.
47. Gerwert K, Liebsch C, Bouton C, Taylor K, Fernández AM, et al. (2023) Diagnostic performance of artificial intelligence and infrared imaging for the histopathologic detection of colorectal cancer. Nat Med 29: 491-502.
48. Whitney J, Corredor G, Janowczyk A, Ganesan S, Doyle S, et al. (2018) Predicting the risk of recurrence in early-stage breast cancer: A pilot study using convolutional neural networks and histopathological images. Am J Pathol 188: 624-633.
49. Wei JW, Suriawinata AA, Vaickus LJ, Ren B, Liu X, et al. (2018) Evaluation of a deep neural network for automated classification of colorectal polyps on histopathologic slides. JAMA Netw Open 1.
50. Shamai G, Geller S, Polak S, Zur H, Segal M, et al. (2019) Artificial intelligence algorithms for breast cancer classification and prediction using histopathological images: A comprehensive review. Cancers (Basel) 11: 1910.
51. Rishi A, Zhang C, Seyedin SN, Hamann S, Kim J, et al. (2020) Predicting molecular subtypes of breast cancer using histopathology images and machine learning. Am J Pathol 190: 819-828.
52. Gamble P, Jaroensri R, Wang H, Durr NJ, Hoffman R, et al. (2021) Stain-independent histopathology using deep learning to interpret unlabelled pathology images. Nat Biomed Eng 5: 432-441.
53. Lagree A, Shiner A, Alera MA, Koohbanani NA, Rajpoot N, et al. (2021) Predicting breast cancer histologic grade from H&E images using convolutional neural networks. Cancer Res 81: 1290-1294.
54. Huff DT, Weisman AJ, Jeraj R (2021) Interpretation and visualization techniques for deep learning models in medical imaging. Phys Med Biol 66: 04TR01.
55. Chen R, Dewi C, Huang S, Caraka RE (2020) Selecting critical features for data classification based on machine learning methods. J Big Data 7: 52.
56. Sakamoto T, Furukawa T, Lami K, Pham HHN, Uegami W, et al. (2020) A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. Translational Lung Cancer Research 9: 2393-2405.
57. Turing A (1950) Computing Machinery and intelligence. Mind 59: 433-460.