AIMS Medical Science

AIMS Medical Science, 12 (4): 350−369.
DOI: 10.3934/medsci.2025024
Received: 23 July 2025
Revised: 30 November 2025
Accepted: 12 December 2025
Published: 19 December 2025

https://www.aimspress.com/journal/medicalScience

Review

Harnessing big data in pathology for precision medicine in gastric cancer: AI-integrated clinical applications

Dongheng Ma[1,2], Hinano Nishikubo[1,2], Tasuku Matsuoka[1,2] and Masakazu Yashiro[1,2,*]

[1] Department of Molecular Oncology and Therapeutics, Osaka Metropolitan University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 5458585, Japan
[2] Cancer Center for Translational Research, Osaka Metropolitan University Graduate School of Medicine, 1-4-3 Asahimachi, Abeno-ku, Osaka 545-8585, Japan

* **Correspondence:** Email: i21496f@omu.ac.jp; Tel: +810666453936.

Abstract: Artificial intelligence (AI) has emerged as a transformative tool in gastric cancer pathology, driving advancements in detection, diagnosis, prognostic modeling, and molecular biomarker identification. Building on these advances, algorithmic innovations such as digital pathology, deep learning, and supervised learning frameworks have facilitated AI integration into clinical practice. Further clinical implementation will require multimodal learning strategies, foundation model development, prospective validation studies, and robust ethical governance. In this review, we provide an updated overview of current applications, technological progress, and prospects for leveraging big data in pathology to achieve AI-driven precision medicine in gastric cancer.

Keywords: gastric cancer; digital pathology; artificial intelligence; deep learning; prognosis prediction; molecular biomarker; whole-slide imaging; cancer pathology

1. Introduction

Gastric cancer (GC) remains a significant global health burden, ranked as the fifth most common malignant cancer and the third leading cause of cancer-related death worldwide [1]. GC shows clinical and histologic heterogeneity [2], posing challenges for tumor diagnosis, particularly in accurate staging and optimal treatment decision-making in clinical practice. Histopathology serves as the diagnostic gold standard [3], providing definitive insights into tumor type, grade, and pattern of invasion.

Digitization of pathology, which generates big data by scanning pathology slides into whole-slide images (WSIs), enables pathologists to analyze the morphologic features of histologic slides using computer-based technology [4].

Artificial intelligence (AI), particularly deep learning (DL), has emerged as a powerful adjunct in pathology, offering high throughput, standardized, and reproducible analyses. DL models applied to digitized pathology slides demonstrate high accuracy in critical clinical tasks, such as tumor detection, diagnostic classification, prognosis prediction, and even molecular biomarker inference [5–7]. Nonetheless, AI-driven precision medicine is still in its infancy and is not yet widely adopted in routine clinical workflows, due to limitations such as model interpretability, generalizability, and data scarcity.

Although AI applications in medical imaging and pathology have been reviewed previously [8,9], there remains a need for an updated, gastric cancer–focused synthesis. In this review, we summarize current AI applications throughout the gastric cancer patient journey, from initial detection and diagnosis to prognostic and molecular prediction. Additionally, we briefly review the technological evolution of AI in digital pathology, discuss current limitations, and outline future directions (Figure 1). We aim to provide clinicians and researchers with a comprehensive overview of the current progress in AI applications in pathology for precision medicine in gastric cancer.

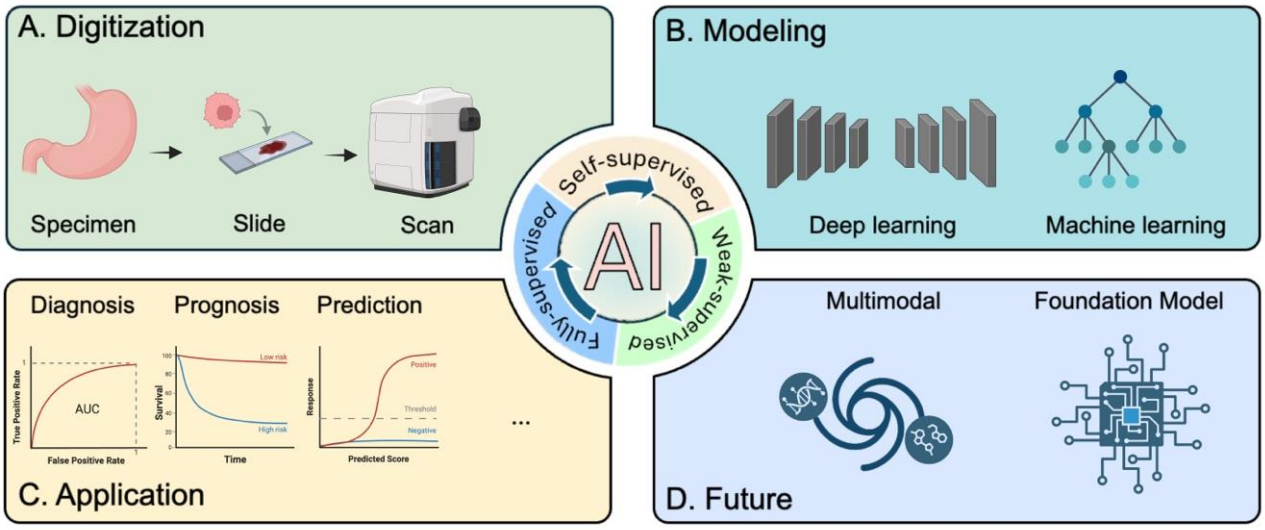

Figure 1. Overview of AI integration in digital pathology for GC. This schematic illustrates the end-to-end workflow of AI in GC pathology. (A) Digitization: Histological analysis begins with the preparation of tissue specimens, after which the slides are scanned into whole-slide images (WSIs) at high resolution. (B) Modeling: DL and machine learning frameworks are trained to perform computational pathology tasks, with supervision strategies evolving from fully supervised to weakly and self-supervised learning. (C) Application: AI models enable a range of applications, including diagnostic classification, prognosis, and treatment response prediction. (D) Future directions: Emerging paradigms such as multimodal integration and foundation models are expected to define future directions. This figure was created with BioRender.com.

1.1. Review methodology

This article is a narrative review. We performed targeted searches of PubMed for English-language studies on artificial intelligence in gastric cancer pathology published from 2020 onward, using combinations of terms such as “gastric cancer”, “digital pathology”, “whole-slide imaging”, “deep learning”, and “computational histopathology”. We also screened the reference lists of key primary studies to identify additional relevant work.

2. Application of pathology AI in GC

Pathology remains a cornerstone of GC management [10], underpinning the full clinical trajectory, including initial detection, diagnosis, and prognosis. However, many of these assessments require labor-intensive, expert-dependent interpretation, which has become more consistent and scalable through AI integration [11]. In this section, we group AI applications according to their position along the gastric cancer care pathway. Detection tasks aim to automatically identify or quantify key pathologic findings such as lymph node metastasis or lymphovascular invasion. Diagnostic tasks focus on histological classification and nodal staging prediction. Prognostic modeling goes one step further to predict outcomes such as overall survival, recurrence, and response to therapy. Molecular prediction tasks aim to infer biomarkers in genomics, transcriptomics, and proteomics from pathology images.

2.1. Detection

The detection of crucial pathologic features is a crucial ground-truth task that underpins downstream clinical decisions [12]. The diagnosis of GC begins with the identification of subtle pathologic features such as tumor infiltration, lympho-vascular invasion, and micro metastasis. Since these features sometimes escape detection by human observers, particularly in high-volume centers, AI-based tools may improve sensitivity and efficiency while minimizing diagnostic oversight.

2.1.1. Lymph node metastasis detection

In gastrectomy specimens, a routine D2 gastrectomy typically yields around 25–40 nodes [13], and micro metastases in slides of lymph nodes (LN) can be easily overlooked [14]. DL might substantially alleviate this burden by segmenting regions with U-Net architectures or highlighting suspicious areas via object detection networks [15]. For instance, Huang et al. [16] trained CNNs using 5-gigapixel images. With 5907 LN images, they achieved an AUC of 0.994 for slide-level metastasis detection, with the micro-metastasis sensitivity rising from 82% (clinical) to 96% (clinical with AI assistance). Their workflow also reduced review time by 31.5% and showed robust performance across multiple centers (AUC 0.983). Similarly, Hu et al. [17], using 921 WSIs, developed an automated system tuned to detect lymph node metastasis, achieving around 0.9 IoU. Beyond detection, AI has enabled the discovery of novel prognostic metrics. For example, a study in *Nature Communications* [18] proposed a DL framework that localized lymph nodes, detected tumor deposits, and computed the tumor-to-lymph node area ratio (T/MLN). This ratio not only correlated strongly with expert assessments but also emerged as an interpretable and independent prognostic factor across multiple cohorts. These studies illustrate how AI significantly improves

detection accuracy, reducing missed diagnoses and supporting pathologists in making more reliable clinical judgments.

2.1.2. Lymph-vascular invasion (LVI) detection

In addition to lymph node metastasis, lymphovascular invasion, wherein tumor cells infiltrate lymphatic or blood vessels, is a critical adverse prognostic marker [19]. AI has shown promising results in detecting LVI. In 2023, Lee et al. [20] combined a transformer-based model (ConViT) with an object detector (YOLOX) to identify LVI foci on H&E-stained slides, achieving an external AUC of 0.94. Another study applied a hard negative mining strategy to train a DL model to detect lymphatic invasion in GC, yielding an AUROC of 0.97 [21]. These advancements demonstrate the feasibility of AI-assisted detection of clinically meaningful yet subtle pathological features. Collectively, these AI-driven detection advances establish a strong foundation for precise diagnosis and prognostic assessment, enhancing clinical decision-making throughout the patient's disease trajectory. However, despite these encouraging results, such models have not yet been widely adopted in routine clinical practice, highlighting the ongoing need for prospective validation and regulatory approval.

2.2. Diagnosis

2.2.1. Diagnosis subtype classification

Accurate histopathological diagnosis remains demanding due to the wide morphologic spectrum of GC [22], from well-differentiated tubular adenocarcinoma to poorly cohesive diffuse types. Recent DL approaches provide reproducible, high-throughput alternatives with expert-level accuracy [23]. In a routine biopsy setting, Park et al. prospectively tested a CNN on 7440 endoscopic biopsies, achieving an AUROC of 0.979 for tumor classification: NFD (no further disease) versus positive (all non-NFD cases), with 100% sensitivity and 97.5% specificity for epithelial lesions. DL assistance reduced average review time by 47%—a compelling case for workload reduction [24].

For general histological classification, Iizuka et al. trained weakly supervised CNN/RNN models on 4128 gastric WSIs, reporting an AUC of 0.980 for classification (adenocarcinoma versus adenoma) on internal data, and the model reached 95.6% accuracy versus 85.9% for 23 pathologists under timed conditions [25]. For subtype classification, Veldhuizen et al. built an attention MIL network that separated Lauren intestinal and diffuse types with a mean AUROC of 0.93. More importantly, they independently stratified 5-year survival in both European and Asian cohorts (HR: 1.4–1.5 vs 1.1–1.2 for pathologists), demonstrating clinical utility beyond morphology [26]. Taken together, these diagnostic advancements underscore how AI provides standardized, reproducible classification systems, which in turn enhance clinical consistency.

2.2.2. Nodal stage prediction

Accurate nodal (N) staging remains pivotal in the diagnosis of GC. Especially for early GC, defined as adenocarcinoma confined to the mucosa (T1a) or submucosa (T1b), patients managed endoscopically (EMR/ESD) do not yield lymph node tissue for histological analysis, and cross-sectional imaging (CT/MRI) is insensitive to micrometastasis. Recent DL models trained directly on

routine H&E slides can now infer metastatic propensity based on primary tumor morphology [27]. Sung et al. [28] were the first to develop an interpretable DL model using morphological features from H&E-stained slides to predict lymph node metastasis (LNM) in early GC, achieving an AUC of up to 0.92 in external validation. For N staging, another study [29] trained an attention-based MIL model on 1146 resections across three hospitals; the lymph node status classifier achieved moderate discrimination (AUROC: 0.71). Guo et al. [30] further integrated multiscale WSI features with baseline clinicopathologic variables; five-fold cross-validation produced a mean AUC of 0.88 for predicting lymph node metastasis on surgical slides and 0.73 on preoperative biopsy slides, demonstrating promising generalizability for real-time surgical decision-making. Collectively, these studies show that slide-based AI can extract nodal metastasis risk signals across the clinical spectrum of GC. In early GC, it provides a decision aid for determining whether to escalate from endoscopic resection to radical surgery or to intensify surveillance. In advanced GC, it offers additional insight to guide surgical planning, determine the extent of lymph node dissection, and potentially stratify patients for adjuvant therapy.

2.3. Prognostic modeling

Beyond refining diagnosis, pathology-based AI has begun to anticipate clinical trajectories directly from routine H&E slides—predicting occult nodal spread, long-term survival, recurrence, and therapeutic response [31].

2.3.1. Predicting OS and recurrence

Conventional pathological prognostic factors—such as TNM stage, tumor differentiation grade, Lauren subtype, and serosal or vascular invasion—remain the backbone of overall survival prediction in GC [32]. Recently, pathology-based DL models have enabled automated and reproducible prognostic stratification based on routine histology slides. Huang et al. [33] developed a multiple instance learning model for gastric cancer (MIL-GC), which was trained on over 1000 patients. Those authors generated a slide-level risk score that remained an independent predictor of overall survival after stage adjustment (HR 1.8) and retained a C-index of 0.657 in external validation, indicating the model correctly ranked the patient with the shorter survival first in 65.7% of comparable pairs. Another study [34] used machine learning to distill 12 quantitative pathomics features into a PSGC signature; when combined with TNM, it increased the 5-year OS AUROC to 0.901 and, crucially, identified stage II–III patients with low PSGC who benefited from adjuvant chemotherapy, whereas high-PSGC patients did not. Moreover, Tian et al. [35] introduced a multiscale DeepRisk network that achieved a C-index of 0.84 for OS. These AI-driven histological features complement conventional staging by offering more granular, reproducible prognostic insights.

Recurrence, particularly occult peritoneal relapse [36], remains the principal driver of postoperative failure in GC. Yet, existing risk models still rely heavily on anatomical factors such as T stage, serosal breach, and nodal ratio [37]. Recent AI-driven studies now show that histology alone can stratify recurrence risk with clinically actionable fidelity. Zhang et al. [38] transformed whole-slide images into compact graphs based on tumor microenvironment features and input them into an adaptive graph clustering network (AGCNet); the model achieved 82% accuracy and an AUC of 0.77 for recurrence prediction in GC. Chen et al. [39] focused on serosa-invasive cases, extracting 186

handcrafted pathomics features from routine H&E tiles and reducing them to an 11-feature signature via LASSO. The signature, when incorporated into a competing risk nomogram alongside CA19-9, depth of invasion, and nodal status, delivered 5-year peritoneal recurrence AUROCs of 0.89 (external validation) and a concordance index of 0.81. These advancements illustrate how AI-derived recurrence features, extracted directly from routine pathology slides, enhance and complement traditional staging systems.

2.3.2. Predicting therapy response

Routine H&E-stained slides, long considered insufficient for predicting treatment response in GC, are now recognized as a valuable substrate for treatment prediction using AI-driven analysis. Recent DL studies [40] show that routine H&E slides alone can stratify patients for immunotherapy, cytotoxic chemotherapy, and targeted therapy. Liu et al. [41] developed ICIsNet, an ensemble model that distilled 148181 biopsy patches into an immune checkpoint inhibitor response score (ICIsRS). Across four independent cohorts, it predicted the benefits of first line PD-1 plus chemotherapy with AUCs of 0.92–1.00 and separated responders from nonresponders at $p < 0.001$. Similarly, another study [42] built a CRSNet from 69564 biopsy patches; it identified pathological major responders to neoadjuvant chemotherapy with an internal AUC of 0.936 and reproduced an AUC of 0.923 externally, again with highly significant score differences between response groups. Zhou et al. [43] applied a contrastive-learning CLAM Sim pipeline to predict *HER2* amplification and downstream trastuzumab efficacy. It reached an AUC of 0.847 for slide-level HER2 status in resections, 0.723 in biopsies, and, crucially, an AUC of 0.833 for distinguishing treatment responders (CR/PR) from nonresponders (SD/PD). These advancements underscore a paradigm shift: AI-based histological biomarkers can serve as universally accessible, cost-effective companion diagnostics, ultimately democratizing precision therapy decisions in GC.

2.4. Molecular prediction

Molecular characterization via DNA mutations, transcriptomic profiles, and protein biomarkers plays a critical role in guiding treatment strategies in GC [44]. However, conventional molecular profiling methods are costly and constrained by limited tissue availability.

2.4.1. Gene mutation

Recent advances have shown promising potential in predicting genetic alterations from H&E-stained slides. For instance, a clinically actionable subtype, Epstein–Barr virus–positive (EBV+) GC, shows strong predictive performance: Vuong et al. developed a DL classifier that predicted EBV status from endoscopic biopsy specimens, reporting an AUC of 0.872 [45]. Microsatellite instability (MSI) and high tumor mutational burden (TMB-H) have also been predicted with moderate to strong success; notably, Hinata et al. reported an MSI prediction model achieving an AUC of 0.880 [46]. However, the broader landscape remains challenging. Attempts to predict mutations in *CDH1*, *ERBB2*, *KRAS*, *PIK3CA*, and *TP53* have yielded more modest AUCs (0.661 to 0.858, Jang et al.) [47], suggesting that some genomic alterations do not produce discernible histopathological features or that existing datasets lack sufficient statistical power. Furthermore, a recent pan-cancer study [48] evaluating 178 GC driver genes found that while 58 achieved a mean AUC ≥ 0.70, only about 5 genes ($\sim$3%) surpassed the more

stringent threshold of mean AUC ≥ 0.80, highlighting the rarity of truly high-performing, histologically predictable mutations in current datasets.

2.4.2. Transcriptomic expression

RNA-level molecular profiling is emerging as a transformative frontier, offering deeper biological insights into tumor heterogeneity and clinical behavior. The initial proof of concept was demonstrated by HE2RNA [49], which predicted gene expression values from 8725 TCGA WSIs spanning 28 tumor types, including 371 GC cases, inferring approximately 3000 genes solely from H&E-stained slides without requiring molecular annotations. Building on this foundation, the linearized transformer model SEQUOIA [50] scaled to 7584 tumors across 16 organs. In the GC subset, it achieved gene-level Pearson correlations of approximately 0.46 and linked predicted pathways, such as inflammatory response and cell cycle regulation, to recurrence risk (Pizurica et al., 2024). The leading edge is now virtual spatial transcriptomics, which employs weak supervision on spatial transcriptomics Visium spots, using H&E slides to infer super-resolution cell spatial profiles that capture cell types, cell states, and their spatial relationships at single-nucleus resolution. Recent AI advancements affirm the clinical viability of RNA inference from pathology images, particularly spatial transcriptomics, establishing a crucial foundation for next-generation, interpretable molecular diagnostics.

2.4.3. Protein biomarkers

Inference of protein expression directly from routine H&E images represents a rapidly evolving field, especially for biomarkers with clear therapeutic implications. Liao et al. [51] used strong supervision to predict HER2 status across a multicenter cohort of 531 WSIs, attaining ~90% external accuracy. Notably, attention heatmaps generated by the model were validated against HER2 immunohistochemistry (IHC) on adjacent tissue slides and demonstrated a strong spatial concordance that strengthens the biological credibility of the model's predictions. Extending beyond oncogene amplification, Jin et al. [52] applied a teacher–student multiple-instance learning framework (MILTS) to 832 stomach adenocarcinoma slides and achieved an AUROC of 0.85 for programmed death ligand 1 (PD-L1). Tile-level probability maps revealed patchy PD-L1-high niches within inflamed glandular mucosa, which showed strong concordance with IHC and made immune checkpoint biology visually accessible to practicing pathologists. For mismatch-repair-deficient (dMMR/MSI-H) tumors, Zheng et al. [53] introduced the ensemble model MMRNet, whose GAN-driven "MMR-Mapping" module achieved an internal AUROC of 0.93 while highlighting syncytial architecture and dense tumor-infiltrating lymphocytes—morphologic hallmarks that enhance the interpretability of otherwise opaque attention heatmaps. These AI-based protein biomarker models deliver not only biomarker-specific predictions but also actionable interpretations that support targeted clinical interventions. Collectively, these molecular prediction capabilities underscore AI's role as a bridge between histopathological images and molecular diagnostics, significantly enriching the clinical utility of pathology throughout the patient journey (Table 1).

Table 1. Summary of studies on protein biomarker prediction in GC.

Study (author)	Year	Model	Training cohort size	Specimen type	Prediction task	Performance
Valieris et al. [54]	2020	ResNet-34	369 patients (TCGA-STAD)	Surgical resections	Gene Signature:MMRD status	AUC: 0.81
Schmauch et al. [49]	2020	HE2RNA	371 patients (TCGA-STAD)	Surgical resections	mRNA expression: 3000~4000 genes	Significant correlation, $P < 0.05$
Kather et al. [55]	2021	ShuffleNet	321 patients (TCGA-STAD)	Surgical resections	Gene mutations: *TP53*, *MTOR*, *FBXW7*, *PIK3CA*; gene expression signatures: proliferation, stemness	Mutation AUC: 0.66~0.78 (gene-dependent); expression signatures AUC > 0.75 (proliferation and stemness)
Jang et al. [47]	2021	CNN-based Model	TCGA-STAD, NS	Surgical resections	Gene mutations: *CDH1*, *ERBB2*, *KRAS*, *PIK3CA*, *TP53*	AUC: 0.661~0.858
Han et al. [56]	2022	CNN-based Model	183 IHC WSIs (Fujian Cancer Hospital, China)	Surgical resections	Molecule: HER2 status	F1: 0.91; accuracy 94%
Jeong et al. [57]	2022	EBVNet	319 patients (TCGA-STAD)	Surgical resections	Gene: EBV status	AUC: 0.88; F1: 0.71
Vuong et al. [45]	2022	EfficientNet	137,184 patches from TMAs and WSIs (Kangbuk Samsung Hospital, Korean)	Surgical resections + biopsies	Gene: EBV status	Biopsy AUC: 0.8723
Flinner et al. [58]	2022	DenseNet161	133 patients (TCGA-STAD)	Surgical resections	Molecular subtypes: EBV, MSI, GS, CIN	AUC: 0.76 (4-class model, external UKC validation)
Lee et al. [59]	2023	CNN-based model	331 patients (TCGA-STAD)	Surgical resections	MSI status	AUC 0.902; AUC 0.968 (external validation)
Wei et al. [60]	2023	CNN-based model	347 patients (TCGA-STAD)	Surgical resections and submucosal dissection	Gene mutations: *KRAS*, *PIK3CA*, *TP53*, *MUC16*	AUC: 0.828~0.923 (external validation)
Jin et al. [52]	2024	MILTS	832 slides (TCGA-STAD)	Surgical resections	mRNA expression: *PD-L1*	AUC: 0.85

Continued on next page

Study (author)	Year	Model	Training cohort size	Specimen type	Prediction task	Performance
Li et al. [61]	2024	CNN-based model	326 patients (TCGA-STAD)	Surgical resections	Gene signature: TMB	AUC: 0.749; multimodal (image + mRNA) AUC: 0.971
Liao et al. [51]	2025	HER2Net	520 patients (Nanfang Hospital, Southern Medical University)	Surgical resections	Molecule: HER2 status	AUC: 0.98
Wu et al. [42]	2025	CLAM_Sim	Surgical: 300 patients; biopsy: 101 patients (Fujian Cancer Hospital, China)	Surgical resections and biopsies	Molecule: HER2 status	AUC: 0.847 (surgical); AUC: 0.723 (biopsy)

*Note: Abbreviations: TMA (tissue microarray), IHC (immunohistochemistry), EBV (Epstein–Barr virus), TMB (tumor mutational burden), MMRD (mismatch-repair deficiency), CLAM (clustering-constrained attention multiple-instance learning), MSI-H (microsatellite instability–high), NS (not specified).

3. Advances in AI algorithms for pathology

3.1. Digital pathology

The transition from conventional glass-slide microscopy to digital pathology marked the foundational step in applying AI to histopathology [62,63]. Early advances in WSI technology enabled the creation of high-resolution, interactive digital slides that preserved the intricate details of histological specimens. Commercial scanners, such as those from Aperio and Hamamatsu, standardized the digitization process and facilitated high-volume data storage and sharing [64]. A regulatory milestone arrived in 2017 when the U.S. FDA granted de novo clearance to the Philips IntelliSite Pathology Solution, authorizing WSI for primary diagnosis [65]. This breakthrough not only legitimized the use of digital slides in routine clinical practice but also unlocked unprecedented opportunities for computational analysis [62]. As hospitals and research institutions began digitizing vast archives of pathology slides, large annotated datasets became available, laying the groundwork for training AI models [66]. Public repositories like the Cancer Genome Atlas (TCGA) and institutional datasets provided diverse, real-world data essential for developing robust, generalizable algorithms [67]. This digital transformation generates massive amounts of big data in pathology, forming the foundation for the development of deep learning algorithms.

3.2. Deep learning emergence

The first generation of computer vision studies in pathology resembled traditional image analysis: investigators hand-crafted features such as texture co-occurrence matrices, color histograms, and Haralick statistics, and fed them into support vector machines or random forests [68]. While these pipelines produced proof-of-concept classifiers, they were brittle; every new stain, scanner, or fixation

protocol necessitated painstaking retuning [69]. Computational pathology researchers pivoted to convolutional neural networks in the mid-2010s; by 2015–2016, CNN–based pipelines were already supplanting hand-engineered feature sets as the default approach [70]. By learning hierarchical representations directly from raw pixels, CNNs rapidly outperformed feature engineering on patch-level cancer detection and subtype classification [71]. U-Net and its medical offspring (Attention-, Res-, and nnU-Net) became the de facto "digital microtome", delivering pixel-accurate gland and tumor bed segmentations that underpinned quantitative histomorphometry and tumor-infiltrating lymphocyte scoring [72]. In GC specifically, where distinguishing epithelial, mucosal, and muscular layers—or annotating particular cell populations such as cancer cells or cancer-associated fibroblasts (CAFs)—is clinically crucial, DL has markedly enhanced both diagnostic support and interpretability [73]. HoVerNet, moreover, introduced a dual-branch architecture that simultaneously separates touching nuclei (via horizontal and vertical distance maps) and labels each nucleus type, enabling studies that correlate spatial immune-cell ecology with prognosis [74]. In parallel, object-detection frameworks such as YOLO eased the manual labeling bottleneck: annotators could mark hundreds of mitoses or signet-ring cells per minute, a capability that accelerated the creation of specialist datasets for GC [75]. Taken together, these developments replaced the ad hoc feature pipelines with end-to-end trainable models, unlocked reliable segmentation of gigapixel slides, and laid the technical groundwork for whole-slide inference, a development that also began to shift the pathologist's role from manual feature identification to AI-assisted decision-making [12].

3.3. Supervised learning

Early DL models demanded that every patch on a whole-slide image be manually labeled—often tens of thousands of marks per case—a process known as "fully supervised learning", which was labor-intensive [76]. The turning point came in 2019 [69], when weak supervision frameworks such as multiple-instance learning (MIL) began treating the entire slide as a single "bag" of tiles and learning, from only a slide-level label, which regions truly matter [77]. The CLAM approach refined this idea by clustering morphologically similar tiles so that the network could distinguish "positive" from "negative" tissue without pixel-level guidance; adapted to GC, it now predicts MSI status and key driver mutations with minimal extra effort from the pathologist [78]. Subsequent MIL variants that introduce transformers or hierarchical pooling follow the same clinical logic: clinicians supply one label per slide, and the algorithm handles the groundwork [79]. In parallel, self-supervised pre-training allows models to learn histologic patterns from millions of unlabeled patches; once fine-tuned with only slide-level labels, these encoders achieve mutation prediction accuracy on par with fully supervised pipelines while remaining robust across scanners and staining protocols [80]. Together, weak and self-supervised learning shift the pathologist's role from laborious annotation to rapid validation, enabling true multicenter studies and bringing AI-driven decision support within reach of everyday gastric cancer care [81].

3.4. Limitations

Despite promising advancements, significant limitations still constrain the practical deployment of AI in GC pathology. First, interpretability remains a major challenge [82]. DL models are often criticized as "black boxes", with thousands of parameters and weights that make it difficult to

understand how decisions are made. While interpretability techniques such as attention-based heatmaps can highlight image regions that influence predictions, they are often coarse and insufficient to support clinical trust or meet regulatory standards. Second, generalizability is limited [83]. Variations in slide scanners, staining protocols, and tissue preparation across institutions and regions can introduce substantial batch effects [84]. As a result, models trained on one dataset may perform poorly on external data [85], and relatively strong internal performance, reflecting overfitting to color distributions, artifacts, or local case-mix rather than robust histological signal, especially when combined with weakly supervised learning that suffers from noisy slide-level labels. Although techniques such as stain normalization, domain adaptation, and generative adversarial networks [86] can mitigate these discrepancies by generating style-invariant virtual slides [87], their success remains variable. Third, data scarcity [88], especially for rare molecular subtypes, poses a serious issue. Certain GC subtypes, such as Scirrhous GC [89], and genomic alterations are underrepresented, resulting in imbalanced training data and reduced predictive power [90]. Strategies such as data collection, oversampling, or basic augmentation (e.g., flipping, rotation) can help increase sample size [91], but they are far from resolving the underlying limitations of data heterogeneity and rarity. Multimodal deep-learning and pathology foundation models can partly mitigate the limitations of data scarcity, but multimodal architecture often involves long, heterogeneous processing pipelines that are difficult to standardize across centers, maintaining limited reproducibility. In summary, improving model transparency, cross-cohort robustness, and rare-case coverage will be crucial for the next phase of AI development in GC pathology.

4. Direction toward clinical implementation

4.1. Multimodal learning

The integration of histopathology with clinical modalities is rapidly reshaping gastric cancer AI. First, the fusion of pathomics and genomics has matured from simple concatenation to true cross-modal reasoning: the M2EF-NNs framework couples Swin Transformer tile embeddings with gene set vectors and, through Dempster–Shafer evidence weighting, improves the C-index/AUC for survival across three TCGA cohorts [92]. Second, pathomics–transcriptomics models such as HE2RNA learn a 256-dimensional "transcriptomic representation" that reconstructs expression of ≈3600 genes per cancer type and, when transferred, boosts MSI detection (AUC 0.81 versus 0.68 on TCGA-CRC) even when no molecular data are available at inference [49]. Third, pathomics–radiopathomic fusion extends these gains to imaging: a Guangxi group extracted 1834 CT radiomics and 512 WSI pathomics features, then combined the top signatures into a nomogram that discriminated stages I–II from III with an external test AUC 0.837, significantly higher than either modality alone [93]. Taken together, these findings show that cross-modal learning enriches H&E morphology, delivering more accurate and biologically interpretable predictions that move AI closer to tangible clinical application.

4.2. Foundation model

Pre-training is rapidly redefining how GC AI models can overcome data scarcity, easing translation into clinical practice. A vision-only foundation model, Virchow, trained on approximately 1.5 million WSIs spanning 17 tissues, delivers a specimen-level AUC of 0.95 across nine common and

seven rare cancers, and still reaches 0.94 on slides from external hospitals while matching, or even surpassing, tissue-specific clinical-grade products despite using far fewer training slides [94]. The visual-language model CONCH pushes label efficiency further: in zero-shot mode, it attains 91.3% accuracy on breast cancer subtyping and, in few-shot experiments, needs only 8 labels per class to outperform models that require 64 [95]. Building on such encoders, the multimodal assistant PathChat fuses a pathology vision backbone with a 13B LLM and achieves 89.5% diagnostic accuracy on 54 multi-organ cases when given images plus clinical context, outperforming GPT-4V and other biomedical chatbots [96]. Altogether, these results show that cross-cancer pre-training, few-shot adaptation, and multimodal alignment are converging into a clinically oriented toolkit. This approach enables rapid fine-tuning on limited gastric cancer data, facilitating direct integration with routine pathological workflows for GC diagnosis.

4.3. Prospective research

A central challenge in the clinical translation of AI for GC lies in the need for high-quality, prospective evidence that demonstrates real-world utility. Several studies have begun to address this gap. Park et al. [24] prospectively evaluated 7440 endoscopic biopsies and showed that a CNN could classify benign versus neoplastic tissue with an AUROC of 0.979 while reducing pathologists' review time by 47% when used as a screening aid. Another study [97] brought AI technology directly into the endoscopy suite: a U-Net-driven femtosecond SRS pipeline produced H&E-like images from fresh biopsies in under 60 s and exceeded 96% diagnostic accuracy, laying the groundwork for real-time intraprocedural decision-making. Complementing these gastric-specific achievements, Raciti et al. [98] delivered the first FDA-cleared evidence that AI-augmented pathology can safely increase sensitivity (+8 pp to 96.6%) and specificity (+0.7 pp) across 610 multicenter prostate biopsies, reducing human detection errors by 70% and proving regulatory and workflow feasibility. Looking ahead, we see that the true driver of clinical translation lies not in incremental performance gains but rather in the accumulation of large-scale, prospective, multicenter, real-world evidence. Such studies not only validate model utility but also foster the clinical confidence essential for integrating AI into patient care.

4.4. Ethical governance

As AI models for digital pathology increasingly rely on patient-derived images and clinical annotations, it will remain crucial to ensure data privacy, mitigate algorithmic bias, and maintain ethical standards [99]. Techniques such as data de-identification and federated learning [100] enable cross-institutional model training without direct sharing, while generative methods such as GANs can produce synthetic yet biologically meaningful images that reduce the risk of patient re-identification [101]. Nevertheless, DL models may still encode latent privacy risks within their learned weights. To address this, many research groups opt to release source code or partially trained models rather than disclosing complete weight files. Because DL models require large, diverse datasets to achieve robust generalization [102], carefully governed data sharing remains crucial to ensure reproducibility and accelerate community-wide progress. At present, most AI systems for GC digital pathology are still used as research tools or decision support, and the final responsibility remains with the pathologist. However, if future workflows were to rely on AI-generated diagnoses, this would

require more prospective validation, allocation of legal responsibility, and explicit patient-informed consent regarding the use of AI tools, given the black-box nature of deep learning models.

5. Conclusions

AI has rapidly advanced the field of GC pathology, offering transformative capabilities across detection, diagnosis, prognostication, and molecular prediction, and paving the way toward precision medicine. As DL methodologies continue to evolve, the integration of multimodal data and the development of pan-cancer foundation models are expected to further expand AI's clinical relevance. Future research should prioritize more interpretable and robust models and studies that evaluate the impact on clinical decision-making. Real-world validation, regulatory approval, and ethical data practices remain critical for translating these advances into routine clinical care. Ongoing collaboration among clinicians, pathologists, and IT scientists is essential to unlock AI's full potential and to realize a future where precision pathology informs every step of GC management.

Author Contributions

Dongheng Ma: methodology, writing—original draft; Hinano Nishikubo: methodology; Tasuku Matsuoka: editing; Masakazu Yashiro: project administration. All authors have read and agreed to the published version of the manuscript.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Part of this study was funded by JST SPRING (Grant Number JPMJSP2139) and KAKENHI (25K11938 to T.M., 24K02525 to M.Y.).

Conflict of interest

The authors declare no conflict of interest.

References

1. Sung H, Ferlay J, Siegel RL, et al. (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 countries. *CA Cancer J Clin* 71: 209–249. https://doi.org/10.3322/caac.21660
2. Matsuoka T, Yashiro M (2024) Molecular mechanism for malignant progression of gastric cancer within the tumor microenvironment. *IJMS* 25: 11735. https://doi.org/10.3390/ijms252111735
3. Zubair M, Owais M, Hassan T, et al. (2025) An interpretable framework for gastric cancer classification using multi-channel attention mechanisms and transfer learning approach on histopathology images. *Sci Rep* 15: 13087. https://doi.org/10.1038/s41598-025-97256-0

4. Jahn SW, Plass M, Moinfar F (2020) Digital pathology: Advantages, limitations and emerging perspectives. *JCM* 9: 3697. https://doi.org/10.3390/jcm9113697
5. Tiwari A, Ghose A, Hasanova M, et al. (2025) The current landscape of artificial intelligence in computational histopathology for cancer diagnosis. *Discov Onc* 16: 438. https://doi.org/10.1007/s12672-025-02212-z
6. Echle A, Rindtorff NT, Brinker TJ, et al. (2021) Deep learning in cancer pathology: A new generation of clinical biomarkers. *Br J Cancer* 124: 686–696. https://doi.org/10.1038/s41416-020-01122-x
7. Unger M, Kather JN (2024) Deep learning in cancer genomics and histopathology. *Genome Med* 16: 44. https://doi.org/10.1186/s13073-024-01315-6
8. Toumaj S, Heidari A, Navimipour NJ (2025) Leveraging explainable artificial intelligence for transparent and trustworthy cancer detection systems. *Artif Intell Med* 169: 103243. https://doi.org/10.1016/j.artmed.2025.103243
9. Farhoudian A, Heidari A, Shahhosseini R (2025) A new era in colorectal cancer: Artificial Intelligence at the forefront. *Comput Biol Med* 196: 110926. https://doi.org/10.1016/j.compbiomed.2025.110926
10. Lordick F, Carneiro F, Cascinu S, et al. (2022) Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol* 33: 1005–1020. https://doi.org/10.1016/j.annonc.2022.07.004
11. Chen S, Ding P, Guo H, et al. (2024) Applications of artificial intelligence in digital pathology for gastric cancer. *Front Oncol* 14: 1437252. https://doi.org/10.3389/fonc.2024.1437252
12. Hosseini MS, Bejnordi BE, Trinh VQ-H, et al. (2024) Computational pathology: A survey review and the way forward. *J Pathol Inform* 15: 100357. https://doi.org/10.1016/j.jpi.2023.100357
13. NCCN (2023) Guidelines for Patients: Stomach Cancer.
14. Tian Y, Pang Y, Yang P-G, et al. (2023) Clinical implications of micro lymph node metastasis for patients with gastric cancer. *BMC Cancer* 23: 536. https://doi.org/10.1186/s12885-023-11023-w
15. Ai S, Li C, Li X, et al. (2021) A State-of-the-Art review for gastric histopathology image analysis approaches and future development. *Biomed Res Int* 2021: 6671417. https://doi.org/10.1155/2021/6671417
16. Huang SC, Chen CC, Lan J, et al. (2022) Deep neural network trained on gigapixel images improves lymph node metastasis detection in clinical settings. *Nat Commun* 13: 3347. https://doi.org/10.1038/s41467-022-30746-1
17. Hu Y, Su F, Dong K, et al. (2021) Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. *Gastric Cancer* 24: 868–877. https://doi.org/10.1007/s10120-021-01158-9
18. Wang X, Chen Y, Gao Y, et al. (2021) Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* 12: 1637. https://doi.org/10.1038/s41467-021-21674-7
19. Japanese Gastric Cancer Association (2023) Japanese Gastric Cancer Treatment Guidelines 2021 (6th edition), *Gastric Cancer* 26: 1–25. https://doi.org/10.1007/s10120-022-01331-8
20. Lee J, Cha S, Kim J, et al. (2024) Ensemble deep learning model to predict lymphovascular invasion in gastric cancer. *Cancers* 16: 430. https://doi.org/10.3390/cancers16020430

21. Lee J, Ahn S, Kim H, et al. (2024) A robust model training strategy using hard negative mining in a weakly labeled dataset for lymphatic invasion in gastric cancer. *J Pathol Clin Res* 10: e355. https://doi.org/10.1002/cjp2.355
22. Shin J, Park YS (2024) Unusual or uncommon histology of gastric cancer. *J Gastric Cancer* 24: 69. https://doi.org/10.5230/jgc.2024.24.e7
23. Zubair M, Owais M, Mahmood T, et al. (2024) Enhanced gastric cancer classification and quantification interpretable framework using digital histopathology images. *Sci Rep* 14: 22533. https://doi.org/10.1038/s41598-024-73823-9
24. Park J, Jang BG, Kim YW, et al. (2021) A prospective validation and observer performance study of a deep learning algorithm for pathologic diagnosis of gastric tumors in endoscopic biopsies. *Clin Cancer Res* 27: 719–728. https://doi.org/10.1158/1078-0432.CCR-20-3159
25. Iizuka O, Kanavati F, Kato K, et al. (2020) Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Sci Rep* 10: 1504. https://doi.org/10.1038/s41598-020-58467-9
26. Veldhuizen GP, Röcken C, Behrens H-M, et al. (2023) Deep learning-based subtyping of gastric cancer histology predicts clinical outcome: a multi-institutional retrospective study. *Gastric Cancer* 26: 708–720. https://doi.org/10.1007/s10120-023-01398-x
27. Budginaite E, Magee DR, Kloft M, et al. (2024) Computational methods for metastasis detection in lymph nodes and characterization of the metastasis-free lymph node microarchitecture: A systematic-narrative hybrid review. *J Pathol Inform* 15: 100367. https://doi.org/10.1016/j.jpi.2024.100367
28. Sung YN (2024) Interpretable deep learning model to predict lymph node metastasis in early gastric cancer using whole slide images. *Am J Cancer Res* 14: 3513–3522. https://doi.org/10.62347/RJBH6076
29. Muti HS, Röcken C, Behrens HM, et al. (2023) Deep learning trained on lymph node status predicts outcome from gastric cancer histopathology: A retrospective multicentric study. *Eur J Cancer* 194: 113335. https://doi.org/10.62347/RJBH6076
30. Guo Z, Lan J, Wang J, et al. (2023) Prediction of lymph node metastasis in primary gastric cancer from pathological images and clinical data by multimodal multiscale deep learning. *Biomed Signal Proces* 86: 105336. https://doi.org/10.1016/j.bspc.2023.105336
31. Srinidhi CL, Ciga O, Martel AL (2021) Deep neural network models for computational histopathology: A survey. *Med Image Anal* 67: 101813. https://doi.org/10.1016/j.media.2020.101813
32. Kulig P, Pach R, Majewska O, et al. (2022) Clinicopathological prognostic factors determining outcomes of treatment in gastric cancer surgery. *In Vivo* 36: 2927–2935. https://doi.org/10.21873/invivo.13035
33. Huang B, Tian S, Zhan N, et al. (2021) Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: A retrospective multicentre study. *EBioMedicine* 73: 103631. https://doi.org/10.1016/j.ebiom.2021.103631
34. Chen D, Fu M, Chi L, et al. (2022) Prognostic and predictive value of a pathomics signature in gastric cancer. *Nat Commun* 13: 6903. https://doi.org/10.1038/s41467-022-34703-w
35. Tian M, Yao Z, Zhou Y, et al. (2024) DeepRisk network: An AI-based tool for digital pathology signature and treatment responsiveness of gastric cancer using whole-slide images. *J Transl Med* 22: 182. https://doi.org/10.1186/s12967-023-04838-5

36. Xiang L, Jin S, Zheng P, et al. (2022) Risk assessment and preventive treatment for peritoneal recurrence following radical resection for gastric cancer. *Front Oncol* 11: 778152. https://doi.org/10.3389/fonc.2021.778152
37. Tonello AS, Capelli G, Bao QR, et al. (2021) A nomogram to predict overall survival and disease-free survival after curative-intent gastrectomy for gastric cancer. *Updates Surg* 73: 1879–1890. https://doi.org/10.1007/s13304-021-01083-7
38. Zhang F, Geng J, Zhang DG, et al. (2023) Prediction of cancer recurrence based on compact graphs of whole slide images. *Comput Biol Med* 167: 107663. https://doi.org/10.1016/j.compbiomed.2023.107663
39. Chen D, Lai J, Cheng J, et al. (2023) Predicting peritoneal recurrence in gastric cancer with serosal invasion using a pathomics nomogram. *iScience* 26: 106246. https://doi.org/10.1016/j.isci.2023.106246
40. Chen Y, Wei K, Liu D, et al. (2021) A machine learning model for predicting a major response to neoadjuvant chemotherapy in advanced gastric cancer. *Front Oncol* 11: 675458. https://doi.org/10.3389/fonc.2021.675458
41. Liu Y, Chen W, Ruan R, et al. (2024) Deep learning based digital pathology for predicting treatment response to first-line PD-1 blockade in advanced gastric cancer. *J Transl Med* 22: 438. https://doi.org/10.1186/s12967-024-05262-z
42. Wu Z, Wang T, Lan J, et al. (2025) Deep learning-based prediction of HER2 status and trastuzumab treatment efficacy of gastric adenocarcinoma based on morphological features. *J Transl Med* 23: 13. https://doi.org/10.1186/s12967-024-06034-5
43. Zhou Z, Ren Y, Zhang Z, et al. (2023) Digital histopathological images of biopsy predict response to neoadjuvant chemotherapy for locally advanced gastric cancer. *Gastric Cancer* 26: 734–742. https://doi.org/10.1007/s10120-023-01407-z
44. Ma D, Fan C, Sano T, et al. (2025) Beyond biomarkers: Machine learning-driven multiomics for personalized medicine in gastric cancer. *JPM* 15: 166. https://doi.org/10.3390/jpm15050166
45. Vuong TTL, Song B, Kwak JT, et al. (2022) Prediction of Epstein-Barr virus status in gastric cancer biopsy specimens using a deep learning algorithm. *JAMA Netw Open* 5: e2236408. https://doi.org/10.1001/jamanetworkopen.2022.36408
46. Hinata M, Ushiku T (2021) Detecting immunotherapy-sensitive subtype in gastric cancer using histologic image-based deep learning. *Sci Rep* 11: 22636. https://doi.org/10.1038/s41598-021-02168-4
47. Jang HJ, Lee A, Kang J, et al. (2021) Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach. *WJG* 27: 7687–7704. https://doi.org/10.3748/wjg.v27.i44.7687
48. Arslan S, Schmidt J, Bass C, et al. (2024) A systematic pan-cancer study on deep learning-based prediction of multi-omic biomarkers from routine pathology images. *Commun Med* 4: 48. https://doi.org/10.1038/s43856-024-00471-5
49. Schmauch B, Romagnoni A, Pronier E, et al. (2020) A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nat Commun* 11: 3877. https://doi.org/10.1038/s41467-020-17678-4
50. Pizurica M, Zheng Y, Carrillo-Perez F, et al. (2024) Digital profiling of gene expression from histology images with linearized attention. *Nat Commun* 15: 9886. https://doi.org/10.1038/s41467-024-54182-5

51. Liao YH, Chen XH, Hu SP, et al. (2025) Artificial intelligence for predicting HER2 status of gastric cancer based on whole-slide histopathology images: A retrospective multicenter study. *Adv Sci* 12: 2408451. https://doi.org/10.1002/advs.202408451
52. Jin D, Liang S, Shmatko A, et al. (2024) Teacher-student collaborated multiple instance learning for pan-cancer PDL1 expression prediction from histopathology slides. *Nat Commun* 15: 3063. https://doi.org/10.1038/s41467-024-46764-0
53. Zheng X, Jing B, Zhao Z, et al. (2024) An interpretable deep learning model for identifying the morphological characteristics of dMMR/MSI-H gastric cancer. *iScience* 27: 109243. https://doi.org/10.1016/j.isci.2024.109243
54. Valieris R, Amaro L, Osório CABDT, et al. (2020) Deep learning predicts underlying features on pathology images with therapeutic relevance for breast and gastric cancer. *Cancers* 12: 3687. https://doi.org/10.3390/cancers12123687
55. Kather JN, Heij LR, Grabsch HI, et al. (2020) Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 1: 789–799. https://doi.org/10.1038/s43018-020-0087-6
56. Han Z, Lan J, Wang T, et al. (2022) A deep learning quantification algorithm for HER2 scoring of gastric cancer. *Front Neurosci* 16: 877229. https://doi.org/10.3389/fnins.2022.877229
57. Jeong Y, Cho CE, Kim JE, et al. (2022) Deep learning model to predict Epstein–Barr virus associated gastric cancer in histology. *Sci Rep* 12: 18466. https://doi.org/10.1038/s41598-022-22731-x
58. Flinner N, Gretser S, Quaas A, et al. (2022) Deep learning based on hematoxylin–eosin staining outperforms immunohistochemistry in predicting molecular subtypes of gastric adenocarcinoma. *J Pathol* 257: 218–226. https://doi.org/10.1002/path.5879
59. Lee SH, Lee Y, Jang H (2023) Deep learning captures selective features for discrimination of microsatellite instability from pathologic tissue slides of gastric cancer. *Int J Cancer* 152: 298–307. https://doi.org/10.1002/ijc.34251
60. Wei Z, Zhao X, Chen J, et al. (2023) Deep Learning–based stratification of gastric cancer patients from hematoxylin and eosin–stained whole slide images by predicting molecular features for immunotherapy response. *Am J Pathol* 193: 1517–1527. https://doi.org/10.1016/j.ajpath.2023.06.004
61. Li J, Liu H, Liu W, et al. (2024) Predicting gastric cancer tumor mutational burden from histopathological images using multimodal deep learning. *Brief Funct Genomics* 23: 228–238. https://doi.org/10.1093/bfgp/elad032
62. Cui M, Zhang DY (2021) Artificial intelligence and computational pathology. *Lab Invest* 101: 412–422. https://doi.org/10.1038/s41374-020-00514-0
63. Kumar N, Gupta R, Gupta S (2020) Whole Slide Imaging (WSI) in pathology: Current perspectives and future directions. *J Digit Imaging* 33: 1034–1040. https://doi.org/10.1007/s10278-020-00351-z
64. Patel A, Balis UGJ, Cheng J, et al. (2021) Contemporary whole slide imaging devices and their applications within the modern pathology department: A selected hardware review. *J Pathol Inform* 12: 50. https://doi.org/10.4103/jpi.jpi_66_21
65. Evans AJ, Bauer TW, Bui MM, et al. (2018) US Food and Drug Administration approval of whole slide imaging for primary diagnosis: A key milestone is reached and new questions are raised. *Arch Pathol Lab Med* 142: 1383–1387. https://doi.org/10.5858/arpa.2017-0496-CP
66. Sakamoto T, Furukawa T, Lami K, et al. (2020) A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. *Transl Lung Cancer Res* 9: 2255–2276. https://doi.org/10.21037/tlcr-20-591

67. Dehkharghanian T, Bidgoli AA, Riasatian A, et al. (2023) Biased data, biased AI: deep networks predict the acquisition site of TCGA images. *Diagn Pathol* 18: 67. https://doi.org/10.1186/s13000-023-01355-3
68. Bankhead P (2022) Developing image analysis methods for digital pathology. *J Pathol* 257: 391–402. https://doi.org/10.1002/path.5921
69. Stacke K, Eilertsen G, Unger J, et al. (2021) Measuring domain shift for deep learning in histopathology. *IEEE J Biomed Health Inform* 25: 325–336. https://doi.org/10.1109/JBHI.2020.3032060
70. Van Der Laak J, Litjens G, Ciompi F (2021) Deep learning in histopathology: the path to the clinic. *Nat Med* 27: 775–784. https://doi.org/10.1038/s41591-021-01343-4
71. Araújo T, Aresta G, Castro E, et al. (2017) Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* 12: e0177544. https://doi.org/10.1371/journal.pone.0177544
72. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation, In: Navab, N., Hornegger, J., Wells, W., Frangi, A. Editors. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. International Publishing: Cham, Springer, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
73. Sharma H, Zerbe N, Klempert I, et al. (2017) Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Comput Med Imaging Graph* 61: 2–13. https://doi.org/10.1016/j.compmedimag.2017.06.001
74. Duanmu H, Wang F, Teodoro G, et al. (2021) Foveal blur-boosted segmentation of nuclei in histopathology images with shape prior knowledge and probability map constraints. *Bioinformatics* 37: 3905–3913. https://doi.org/10.1093/bioinformatics/btab418
75. Debsarkar SS, Aronow B, Prasath VBS (2025) Advancements in automated nuclei segmentation for histopathology using you only look once-driven approaches: A systematic review. *Comput Biol Med* 190: 110072. https://doi.org/10.1016/j.compbiomed.2025.110072
76. Xu H, Usuyama N, Bagga J, et al. (2024) A whole-slide foundation model for digital pathology from real-world data. *Nature* 630: 181–188. https://doi.org/10.1038/s41586-024-07441-w
77. Gadermayr M, Tschuchnig M (2024) Multiple instance learning for digital pathology: A review of the state-of-the-art, limitations & future potential. *Comput Med Imaging Graph* 112: 102337. https://doi.org/10.1016/j.compmedimag.2024.102337
78. Wang Z, Peng H, Wan J, et al. (2024) Identification of histopathological classification and establishment of prognostic indicators of gastric adenocarcinoma based on deep learning algorithm. *Med Mol Morphol* 57: 286–298. https://doi.org/10.1007/s00795-024-00399-8
79. Chen RJ, Chen C, Li Y, et al. (2022) Scaling vision transformers to gigapixel images via hierarchical self-supervised learning, *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 16123–16134. https://doi.org/10.1109/CVPR52688.2022.01567
80. Saldanha OL, Loeffler CML, Niehues JM, et al. (2023) Self-supervised attention-based deep learning for pan-cancer mutation prediction from histopathology. *NPJ Precis Oncol* 7: 35. https://doi.org/10.1038/s41698-023-00365-0
81. Shmatko A, Ghaffari Laleh N, Gerstung M, et al. (2022) Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat Cancer* 3: 1026–1038. https://doi.org/10.1038/s43018-022-00436-4

82. Faa G, Fraschini M, Barberini L (2024) Reproducibility and explainability in digital pathology: The need to make black-box artificial intelligence systems more transparent. *J Public Health Res* 13: 22799036241284898. https://doi.org/10.1177/22799036241284898
83. Lin S, Zhou H, Watson M, et al. (2025) Impact of stain variation and color normalization for prognostic predictions in pathology. *Sci Rep* 15: 2369. https://doi.org/10.1038/s41598-024-83267-w
84. Nunes JD, Montezuma D, Oliveira D, et al. (2025) Bridging domain gaps in computational pathology: A comparative study of adaptation strategies. *Sensors* 25: 2856. https://doi.org/10.3390/s25092856
85. Howard FM, Dolezal J, Kochanny S, et al. (2021) The impact of site-specific digital histology signatures on deep learning model accuracy and bias. *Nat Commun* 12: 4423. https://doi.org/10.1038/s41467-021-24698-1
86. Breen J, Zucker K, Allen K, et al. (2024) Generative adversarial networks for stain normalisation in histopathology, In: Lyu, Z. Editor. *Applications of Generative AI*, Netherland: Springer, Cham, 227–247. https://doi.org/10.1007/978-3-031-46238-2_11
87. Voon W, Hum YC, Tee YK, et al. (2023) Evaluating the effectiveness of stain normalization techniques in automated grading of invasive ductal carcinoma histopathological images. *Sci Rep* 13: 20518. https://doi.org/10.1038/s41598-023-46619-6
88. Tran KA, Kondrashova O, Bradley A, et al. (2021) Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* 13: 152. https://doi.org/10.1186/s13073-021-00968-x
89. Ayub A, Naeem B, Perez A, et al. (2023) Gastric linitis plastica: clinical characteristics and outcomes from the national cancer database. *Anticancer Res* 43: 1543–1548. https://doi.org/10.21873/anticanres.16303
90. Unger M, Kather JN (2024) A systematic analysis of deep learning in genomics and histopathology for precision oncology. *BMC Med Genomics* 17: 48. https://doi.org/10.1186/s12920-024-01796-9
91. Faryna K, Van Der Laak J, Litjens G (2024) Automatic data augmentation to improve generalization of deep learning in H&E stained histopathology. *Comput Biol Med* 170: 108018. https://doi.org/10.1016/j.compbiomed.2024.108018
92. Luo H, Huang J, Ju H, et al. (2025) Multimodal multi-instance evidence fusion neural networks for cancer survival prediction. *Sci Rep* 15: 10470. https://doi.org/10.1038/s41598-025-93770-3
93. Tan Y, Feng L, Huang Y, et al. (2024) A comprehensive radiopathological nomogram for the prediction of pathological staging in gastric cancer using CT-derived and WSI-based features. *Transl Oncol* 40: 101864. https://doi.org/10.1016/j.tranon.2023.101864
94. Vorontsov E, Bozkurt A, Casson A, et al. (2024) A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat Med* 30: 2924–2935. https://doi.org/10.1038/s41591-024-03141-0
95. Lu MY, Chen B, Williamson DFK, et al. (2024) A visual-language foundation model for computational pathology. *Nat Med* 30: 863–874. https://doi.org/10.1038/s41591-024-02856-4
96. Lu MY, Chen B, Williamson DFK, et al. (2024) A multimodal generative AI copilot for human pathology. *Nature* 634: 466–473. https://doi.org/10.1038/s41586-024-07618-3
97. Liu ZJ, Su W, Ao JP, et al. (2022) Instant diagnosis of gastroscopic biopsy via deep-learned single-shot femtosecond stimulated Raman histology. *Nat Commun* 13: 4050. https://doi.org/10.1038/s41467-022-31339-8

98. Raciti P, Sue J, Retamero JA, et al. (2023) Clinical validation of artificial intelligence–augmented pathology diagnosis demonstrates significant gains in diagnostic accuracy in prostate cancer detection. *Arch Pathol Lab Med* 147: 1178–1185. https://doi.org/10.5858/arpa.2022-0066-OA
99. Ziller A, Mueller TT, Stieger S, et al. (2024) Reconciling privacy and accuracy in AI for medical imaging. *Nat Mach Intell* 6: 764–774. https://doi.org/10.1038/s42256-024-00858-y
100. Lu MY, Chen RJ, Kong D, et al. (2022) Federated learning for computational pathology on gigapixel whole slide images. *Med Image Anal* 76: 102298. https://doi.org/10.1016/j.media.2021.102298
101. Chen RJ, Lu MY, Chen TY, et al. (2021) Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 5: 493–497. https://doi.org/10.1038/s41551-021-00751-8
102. Song AH, Jaume G, Williamson DFK, et al. (2023) Artificial intelligence for digital and computational pathology. *Nat Rev Bioeng* 1: 930–949. https://doi.org/10.1038/s44222-023-00096-8

AIMS Press