Discover Oncology

Review

The current landscape of artificial intelligence in computational histopathology for cancer diagnosis

Aaditya Tiwari[1,2,3] · **Aruni Ghose**[2,3,4,5,6,7,18,19] · **Maryam Hasanova**[7,8] · **Sara Socorro Faria**[9] · **Srishti Mohapatra**[10,11] · **Sola Adeleke**[12,13] · **Stergios Boussios**[5,14,15,16,17]

Received: 15 December 2024 / Accepted: 24 March 2025
Published online: 01 April 2025

Abstract
Artificial intelligence (AI) marks a frontier in histopathologic analysis shift towards the clinic, becoming a mainstream choice to interpret histological images. Surveying studies assessing AI applications in histopathology from 2013 to 2024, we review key methods (including supervised, unsupervised, weakly supervised and transfer learning) in deep learning-based pattern recognition in computational histopathology for diagnostic and prognostic purposes. Deep learning methods also showed utility in identifying a wide range of genetic mutations and standard pathology biomarkers from routine histology. This survey of 41 primary studies also encompasses key regions of AI applicability in histopathology in a multi-cancer review while marking prospects to introduce AI into the clinical setting with key examples including Swarm Learning and Data Fusion.

Keywords Artificial Intelligence · Machine Learning · Deep Learning · Cancer · Histopathology · Computational

1 Introduction

In 2022, there were 20 million new cancer cases and 10 million cancer-related deaths recorded. The commonest incident tumour types were lung (12.4%), breast (11.6%), and colorectal (9.6%) while mortality was highest for lung (18.7%), colorectal (9.3%) and liver (7.8%) [1]. Accurate, rapid detection minimises mortality and morbidity. Identifying cancer type, tumour markers, and invasion is vital in precision oncology [2]. Biopsy remains the gold standard for confirming malignancy. In the modern era, the mainstay of histopathological analysis is digitised via whole slide imaging (WSI). Digitised histopathological analysis via whole-slide imaging (WSI) enables computational histopathology to extract

✉ Aruni Ghose, aruni.ghose1@gmail.com | [1]Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [2]Barts Cancer Institute, Cancer Research UK City of London Centre, Queen Mary University of London, London, UK. [3]Department of Oncology, Princess Alexandra Hospital NHS Trust, Harlow, UK. [4]Barts Cancer Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, UK. [5]Department of Medical Oncology, Medway NHS Foundation Trust, Gillingham, UK. [6]Digital Health Network, European Cancer Organisation, Brussels, Belgium. [7]OncoFlowTM, London, UK. [8]Division of Biosciences, University College London, London, UK. [9]Laboratory of Immunology and Inflammation, Department of Cell Biology, University of Brasilia, Brasilia, DF, Brazil. [10]General Internal Medicine Doctorate Programme, University of Hertfordshire, Hatfield, UK. [11]The Misdiagnosis Association and Research Institute, California, USA. [12]School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK. [13]Guy's Cancer Centre, Guy's and St. Thomas' NHS Foundation Trust, London, UK. [14]Kent and Medway Medical School, University of Kent, Canterbury, UK. [15]Faculty of Medicine, Health and Social Care, Canterbury Christ Church University, Canterbury, UK. [16]Faculty of Life Sciences and Medicine, School of Cancer & Pharmaceutical Sciences, King's College London, London, UK. [17]AELIA Organization, 9Th Km Thessaloniki-Thermi, 57001 Thessaloniki, Greece. [18]United Kingdom and Ireland Global Cancer Network, Manchester, UK. [19]Oncology Council, Royal Society of Medicine, London, UK.

clinically relevant information using algorithmic models. Advances in Big Data, Machine Learning (ML), and AI hold immense potential for rapid cancer diagnosis [3].

Computational histopathology integrates raw data including medical records, "-omics" data, and radiological/pathological imaging to highlight biologically and clinically relevant information [4]. Omics"(genomics, epigenomics, transcriptomics, proteomics, and metabolomics) demonstrates increasing value, with AI essential for multi-omics data integration. Computational histopathology broadens the scope by characterising multi-omics aberrations via histopathologic images, predicting prognosis [5]. Sophisticated AI mechanisms maximise diagnostic potential, enabling time-efficient cancer characterisation, early detection, prompt treatment, and improved prognosis.

ML uses data and algorithms to learn from large datasets, self-improving to yield readable information without explicit programming (Fig. 1) [6]. The multifaceted nature of healthcare problems suits ML, implementing statistical and mathematical optimisation to exploit patterns and adapt output (Table 1) [7]. These introduce objectivity and remove interpersonal variability which often limits traditional histopathological analysis. Such approaches have already shown success in radiology [8, 9], endoscopy [10], and dermoscopy [11].

ML in computational histopathology is a recent development involving high-quality WSIs and progress in image development at reduced costs. AI in computational histopathology has shown exponential improvement in accuracy, with error rates in the classification of natural images decreasing from 25 to 4% in three years [12]. Deep learning (DL) methods have broad applications in cancer detection, quantification, and typing, from rudimentary single small-area analysis [13] to breast cancer (BC) segmentation [14], and glioma classification [15].

This article describes basic ML techniques, comprehensively assessing and critiquing their applications in computational histopathology in a multi-cancer survey to address current utilities and future outlooks.

2 AI for quantifying tumour extent

2.1 Supervised models

Supervised learning requires pre-training, with each data point possessing a corresponding output label or ground truth. This enables the algorithm to learn input–output relationships, allowing predictions on new, unlabelled data. This has utility in histopathological image classification for predictions based on visual features (Table 2).

Image classification such as detecting cells or nuclei is among the most successful tasks with DL techniques significantly advancing digital pathology. Ciresan et al. developed convolutional neural network (CNN)-based pixel prediction for mitosis detection in H&E BC histology images [13]. Wang et al. formulated a cascade ensemble of CNN and handcrafted features for mitosis detection [16]. These methods utilise pixel-wise prediction tasks by sliding windows training networks on small image patches rather than entire WSIs. Trained pathologists annotate the image patches containing objects of interest (cells/nuclei). However, mitotic figures are often rare events creating class imbalance issues. CNNs trained on

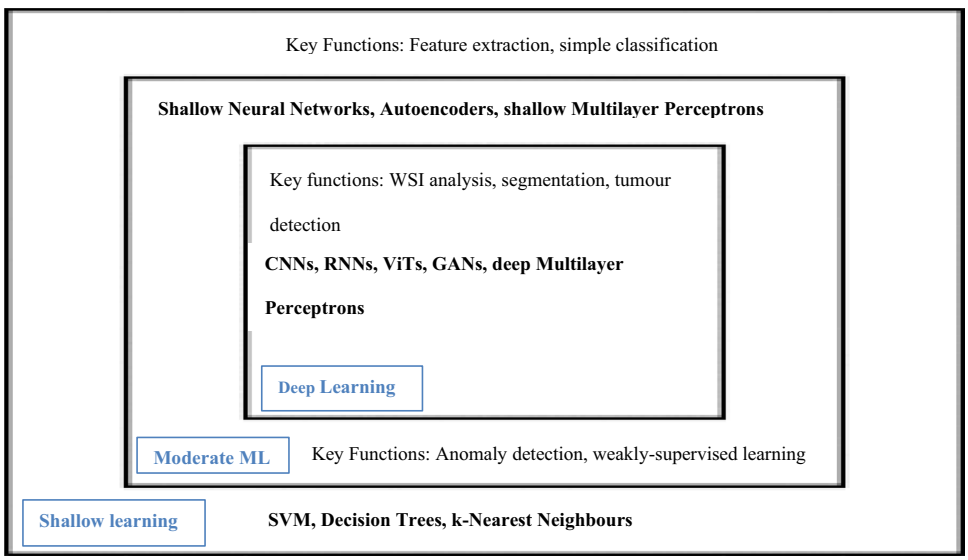

Fig. 1 diagram of ML concepts with specific examples

Table 1 Selected methods of ML, basic functionality, and application in computational histopathology

Machine learning type	Description	Example algorithms	Key application in histopathology
Supervised Learning	Labelled data with known input–output mapping	CNNs, Random Forest, SVM, ResNet	Tissue segmentation, cancer classification, and biomarker identification
Weakly-supervised Learning	Coarse/incomplete labels at slide level instead of pixel level	MIL, CAM-based models, WS-DAN	WSI classification, tumour localisation
Unsupervised Learning	Identify patterns without labelled data and predefined outputs	K-means clustering, DBSCAN, PCA, Autoencoders	Cell clustering, anomaly detection, and feature extraction
Transfer Learning	Adapt pre-trained models to histopathology tasks	ResNet, DenseNet	Reduced training time for tissue segmentation, cancer classification, and anomaly detection
Self-supervised Learning	Pseudo-label generated from data for meaningful representations	SimCLR, MoCo, BYOL, MAE	Pretraining for histopathology images and feature extraction
Vision Transformers	Attention mechanisms for image analysis	ViT, Swin Transformer, DeiT	WSI classification, feature extraction, and long-range dependency modelling

Table 2 Key examples of supervised learning methods in histopathology summarised

Study	Cancer types	Stain	Application	Method	Dataset
Ciresan et al. [13]	Breast cancer	H&E stain	Detection of mitotic figures in breast cancer	CNN-based pixel prediction	50 images corresponding to 50 high-power fields partitioned into three subsets (T1 = 26, T2 = 9, T3 = 15)
Ertosun et al. [15]	Gliomas	H&E stain	Glioma grade classification	CNN model segmenting nuclei for cancer localisation and patch-based glioma classification	875 e-micro-biopsy samples from 22 WSIs from 4 source sites. Training subset = 6998. Validation subset = 1752
Sirinukunwattana et al. [17]	Colorectal adenocarcinoma	H&E stain	Cell and nucleus detection without nuclei segmentation	SC-CNN coupled with Neighbouring Ensemble Predictor	> 20,000 annotated nuclei in four different classes: epithelial, inflammatory, fibroblast, miscellaneous
Xu et al. [29]	Colorectal adenocarcinoma	H&E stain	Gland segmentation and instance recognition	Multi-task network cascades with CNN and Simultaneous Detection and Segmentation	165 annotated histological images: 85 images in training set and 80 images in test set
Zhang et al. [25]	Urothelioma	H&E stain	Bladder cancer diagnosis augmentation	CNN for tumour detection and cellular-level characterization and RNNs to control visual and linguistic competencies	913 WSIs obtained from TCGA (open sourced)
Graham et al. [26]	Colorectal adenocarcinoma	H&E stain	Automated nuclear segmentation and classification	Novel CNN for simultaneous segmentation and classification for large-scale nuclear morphometry	24,319 annotated nuclei with associated class labels
Gong et al. [33]	Breast cancer	H&E stain	Intrinsic feature representation	Self-distilled supervised contrastive learning for CNN -based CAD with limited training samples	4 breast cancer datasets: 78 normal tissue, 72 benign lesions, 71 in situ, 64 invasive carcinomas)
Abousamra et al. [34]	Multiple cancers	H&E stain	Patch classification as TIL (tumour-infiltrating lymphocyte) positive or negative	VGC-16, ResNet-34, Inception-V4 models trained on ImageNet to output value per patch in input WSI	7983 WSIs across 23 cancers retrieved from TCGA
Sarker et al. [35]	Breast cancer	H&E stain	Breast cancer classification	FMB-ConV and MBConv with DSE for categorising into binary (benign/malignant) + eight subtype classification	Publicly available BreakHis dataset—7909 images at 40x, 100x, 200x, 400 × magnifications

imbalanced datasets may favour the majority, non-mitotic class. Additionally, histopathological images exhibit variability and artefacts from staining and slide preparation methods reducing generalizability and output quality (Tables 3, 4).

Sirinukunwattana et al. used spatially constrained CNN (SC-CNN) for nuclei detection utilising hand-crafted features on 100 H&E stained colorectal adenocarcinoma images with 29,756 nuclei manually annotated by a pathologist [17]. This CNN analysed texture features demonstrating that CNNs were effective in automatically learning features: however, better performance was achieved when combining manually handcrafted features with raw data. Additional SC-CNN features included a colour deconvolution method to separate the haematoxylin channel [18] and a scattering coefficient to add texture information providing stable, invariant descriptors for classification [19].

Xu et al. employed a cutting-edge stacked sparse autoencoder (SSAE) algorithm for nucleus detection [20]. A radius of 6 pixels was considered the ground truth at the annotated epithelial nuclei centre. Any detection point within this area was the true positive. If multiple detection points were within the same ground truth area, only one was counted as positive. Figure 2 shows probability mapping results comparing SC-CNN and SSAE: detected results are shown as green dots against the yellow "ground truth". Proximity mapping identified the maxima. SC-CNN missed some detected cells (highlighted in red). The detection accuracy of the deep CNN was improved by handcrafted features. Different quantitative performance measures could be effective in capturing tumour nuclei appearance which could upscale other cells (e.g., immune/non-cancerous).

Ertosun et al. used CNN-based molecular classification frameworks using The Cancer Genome Atlas (TCGA) WSIs to grade gliomas from digital pathology images [15]. The dataset included Lower-Grade Glioma (LGG) Grades II and III, and Glioblastoma Multiforme (GBM) Grade IV [21]. TCGA WSIs yielded data structures cropped into "tiles or patches", for parallel processing and to accelerate pre-processing. Following nuclear morphology-based studies in GBM [22, 23], nuclei were segmented during the pre-processing stage, but the individual nuclei were not cropped. Images were resized to scales suitable for DL on the graphics processing unit (GPU). However, the CNNs were not cost-effective for high-resolution images and had limited memory availability. Image size was adjusted to 256×256, considering memory capacity and time to train neural networks.

Ertosun et al. noted that LGG grade classification had a 71% accuracy when trained on independent data [15]. This performance was limited due to preparative protocol variations such as differences in microscopes, digitisation, and slide preparation; collectively called *"batch effects"*. WSI quality was thus marred by artefacts. GBM cases showed clear distinction due to their low survival probability. LGG Grade II and Grade III showed similarity with a 95% confidence interval (CI) overlapping separate survival curves. CNN modules performed similarly to observations from other studies demonstrating challenges in grading intermediate tumours with a strong survival overlap [24]. Although pathologists can distinguish LGG and GBM with similar accuracy, this study provides a valuable *"second look"* or teaching application.

Results could improve through a modular approach to optimise task components like GBM vs LGG diagnosis and determination of tumour grade separately. Global CNN architecture was used for tissue-level cancer localization and WSI-level disease mostly for patch-based classification. This enhanced diagnostic quality statistics, simplifying train models despite limited data availability [15].

Zhang et al. used CNN and Recurrent Neural Networks (RNN) to generate clinical diagnostic descriptions and visual attention maps for urothelial cancer diagnosis, from a dataset of 913 images. Their novel WSI method translated gigapixels into a series of interpretable predictions, providing second opinions and demonstrating clinical applications. Their DL model matched the effectiveness of 17 pathologists, advancing low-cost, next-generation AI-enhanced diagnostics in pathology [25]. Graham took this further, proposing a unified FCN model to simultaneously identify nuclei in both horizontal and vertical planes of pixels to determine nuclei separation in a 3-D manner in multi-tissue histology images [26].

Regression models with DL in histopathology map image patches to identify nuclear positions and estimate density maps to compare normal and tumour topologies. Graham et al. effectively applied these via CNN-based nuclei segmentation and classification across multiple cancers [26]. Their evaluation framework quantified nuclear segmentation using a new dataset of 24,319 annotated nuclei with class labels. Their model detected nuclear pixels and a post-processing pipeline simultaneously segmented nuclear instances to obtain corresponding nuclear types. The down-sampling factor was reduced from 32 to 8 using a stride size of 1 in the first convolution and removing max-pooling. The spatial dimension of the input was 270×270 pixels, with 80×80 output dimensions. Deep neural networks (DNNs) used feature extraction, inspired by pre-trained residual networks with ResNet-50 layers [27] ensuring robustness against input disruption [28].

Graham et al. evaluated instance segmentation performance using five datasets (CoNSeP, CPM-15, Kumar, CPM-17, and TNBC). A major challenge was limited useful data with associated class labels. For instance, CoNSeP contained over 24,000 labelled nuclei from difficult samples highlighting the complexity of segmenting nuclei in WSIs despite abundant data. True nuclear segmentation performance was assessed but measurements often failed to reflect actual instance

Table 3 Key examples of weakly-supervised learning methods in histopathology summarised

Study	Cancer types	Stain	Application	Method	Dataset
Hou et al. [36]	Brain	H&E stain	Glioma subtype classification	Expectation–maximisation-based MIL with CNN + logistic regression	TCGA (1064 slides)
Jia et al. [38]	Colon	H&E stain	Segmentation of cancerous regions	FCN-based MIL + deep supervision and area constraints	Two private sets containing colon cancer images (910 + 60)
Liang et al. [39]	Stomach	H&E stain	Gastric tumour segmentation	Patch-based FCN + iterative learning approach	China Big Data and AI Challenge (1900 images)
Wang et al. [40]	Stomach	H&E stain	Gastric cancer detection	Two-stage CNN framework for localization and classification	Private set (608 images)
Li et al. [45]	Breast	H&E stain	CNN + multi-branch attention modules and deep supervision mechanism	CNN + multi-branch attention modules and deep supervision mechanism	PCam (327,680 patches extracted from Camelyon16) and Camelyon-16 (400 WSIs)
Wang et al. [43]	Lung	H&E stain	Lung cancer image classification	Patch-based FCN + context-aware block selection and feature aggregation strategy	Private (939 WSIs), TCGA (500 WSIs)
Campanella et al. [42]	Multi-cancers	H&E stain	Multiple cancer diagnoses in WSIs	CNN (ResNet) + RNNs	Prostate (24,859 slides), skin (9,962 slides), breast cancer metastasis (9,894 slides)
Guo et al. [48]	–	–	Cell nuclei segmentation in histopathology images	Weakly-supervised learning SAC-Net trained in pseudo ground truth labels	21,623 nuclei MoNuSeg, 4022 nuclei TNBC, CPM-17 7570 nuclei
Zheng et al. [49]	Renal cell carcinoma	H&E stain	Human–machine fusion for categorization of clear cell renal cell carcinoma	SSL-CLAM for automatic detection of subregions with high diagnostic value	445 WSIs

Table 4 Key examples of unsupervised learning methods in histopathology summarised

Study	Cancer types	Stain	Application	Method	Dataset
Hu et al. [20]	Breast cancer	H&E stain	Nuclei detection on high-resolution histopathological images	SSAE for nuclei detection categorising images as nuclear/non-nuclear	500 histopathological images and approximately 3500 manually segmented individual nuclei
Bulten et al. [54]	Prostate cancer	H&E stain	Classification of PC as tumour/non-tumour without prior training	Self-clustering convolutional adversarial autoencoder to learn relevant features for the clustering task	94 registered WSI pairs into training (54) and test (40) in patients that underwent radical prostatectomy
Hu et al. [53]	Blood disease	H&E stain	Unsupervised classification of bone marrow cellular components	Cell-level visual representation using GAN for nuclei segmentation	Dataset A: publicly available – eleven healthy images Dataset B: WSIs of bone marrow from 24 patients with blood diseases Dataset C: Database A + B Database D: WSIs from 28 patients with bone marrow haematopoietic tissue hyperplasia + 56 patients with leukaemia
Sari et al. [55]	Colon	H&E stain	Representation and classification of histopathological tissue	Quantization of salient subregions by feature extraction	Dataset 1 – 3236 images from 258 patients Dataset 2 – derived from dataset 1 with further low-grade tissue subcategorization
Sheikh et al. [56]	Multi-cancers	H&E stain	Generalise different cellular heterogeneity representations from multi-image descriptors	Unsupervised learning with oriented gradients, local binary patterns, fusion of heterogeneous features	3800 samples of invasive class WSIs
Nam et al. [57]	Lung adenocarcinoma	–	Identify patients with histopathologic risk factors via CT-based DL	Unsupervised clustering and regression analyses to estimate disease-free survival in patients with resected lung adenocarcinoma	1667 patients with resected, pathological stage I-IV lung adenocarcinoma
Niehues [58]	Colorectal cancer	H&E stain	Biomarker identification from WSIs	Attention-based SSL to combine global features and minimise artefact interference	PCam (327,680 patches extracted from Camelyon16) and Camelyon 16 (400 WSIs)
Hosseini et al. [59]	Prostate cancer	H&E stain	Classification of prostate cancer into low/high cancerous regions	Label assignment to unlabelled data points to select high compatibility targets	PANDA*/Karolinska – 5414 images PANDA/Radboud – 5160 images TMA1**/Gleason 2019 – 244 TMA2/Harvard—641
Fetisov et al. [60]	Prostate cancer	–	Segmentation of histopathology images in prostate cancer to self-determine important regions of images	Extracted feature vectors used for K-means clustering and self-training to produce new accuracies and compute feature scores	PANDA dataset with 10,616 cases and one image from each case

Table 4 (continued)

Study	Cancer types	Stain	Application	Method	Dataset
Wang et al. [61]	Multi-cancers	H&E	Patch retrieval, classification, mitosis detection and colorectal adenocarcinoma gland segmentation	SRCL integrated with hybrid CTransPath	15 million patches from 30,000 WSIs spanning 11 datasets (TCGA, PAIP, UniToPatho, TissueNet, NCT-CRC-HE, CRC, Camelyon16, TCGA-NSCLC, TCGA-RCC, MIDOG, CRAG)
Dippet et al. [62]	Multi-pathology	Multiple stain	Semantic slide and tissue patch grouping	DINOv2 framework with integrated pathologist domain knowledge	103,849 WSIs

*TMA—Tissue Microarray, **PANDA—Prostate cANcer graDe Assessment

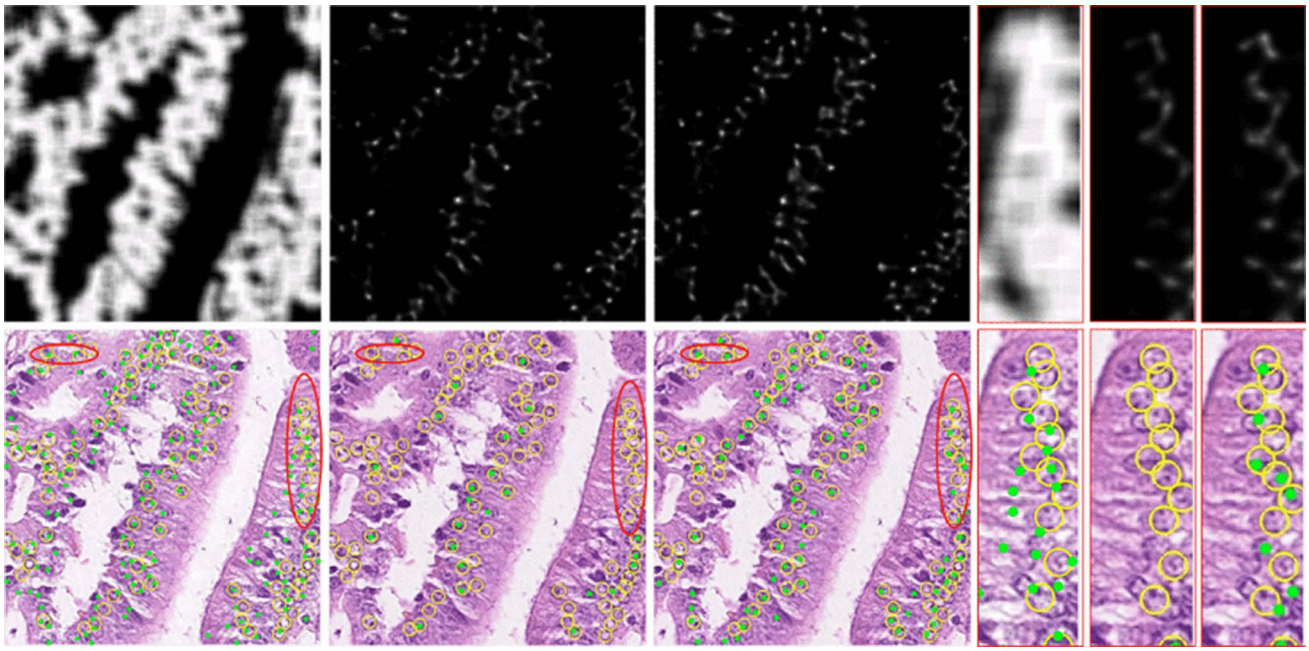

Fig. 2 Column 1 – proximity map and detection row using SSAE. Column 2 – proximity map and detection with SC-CNN. Column 3 – proposed approach. Columns 4–6 – Zoomed in image of Columns 1–3 respectively [20]

segmentation performance. Other methods explored relationships of nuclear types in histology images but were limited to spatial analysis unavailable. segmentation. Despite challenges, the model did provide opportunities to explore spatial relationships [26].

Xu et al. employed multi-channel DNN models for gland segmentation and instance recognition using the GLaS challenge dataset in colon cancer consisting of 165 images [29]. Their image-to-image prediction fused multi-layered information including regional, locational, and boundary cues in histological images. Their model addressed complex feature design issues with CNNs altering channels to meet multi-level requirements. Instance segmentation enabled detailed image processing by distinguishing contour, location, class, and object count [29]. Traditional algorithms struggle with complex problems, demonstrating demand for DL. Simultaneous Detection and Segmentation (SDS) proposes a framework to solve gland instance segmentation, hyper-column, and MNC (Multi-task Network Cascades) optimising and accelerating the feature extraction process. Xu et al. MNC framework optimised gland segmentation and gland instance simultaneously with a three-channel algorithm, each with different responsibilities.

Xu et al. [29] highlighted limitations in gland instance segmentation. The lack of gland-separating tissues makes segmentation challenging: the algorithm may misidentify a single connecting pixel as part of the gland structure. Secondly, the algorithm's separate instance recognition can lead to smaller prediction areas than the ground truth, reducing segmentation performance despite accurate number and position (Fig. 3) [29]. Multichannel neural networks proved equal effectiveness as methods reported in the 2015 MICCAI Gland Segmentation Challenge [30] implying that the model's generalisation ability enabled the algorithm to solve gland instance segmentation problems and also its adaptability for specific tasks. Baseline experiments confirmed superiority for instance segmentation.

Histological objects e.g. glands and nuclei, vary significantly based on size, shape, and colour with overlapping clumped areas complicating distinction from surrounding structures.. Beyond Xu et al. [29], other studies used Fully Convolutional Networks (FCN) to distinguish surrounding structures, leveraging multi-scale feature learning to address histology image variation [31]. Conventional FCN-based models are fundamentally designed to predict class labels as foreground or background but cannot identify individual object instances [32].

Gong et al. applied Self-Distilled Supervised Contrastive Learning (SDSCL) for CNN-based computer-aided diagnosis (CAD) on H&E images to enhance intrinsic feature representation [33]. Combining supervised and contrastive learning, they trained a CNN to self-distil, addressing limited sample availability. Four datasets (78 normal tissue, 72 benign lesions, 71 in-situ carcinoma, and 64 invasive carcinoma) were digitised and annotated by two pathologists. They utilised ResNet-18 pre-trained on ImageNet weights and a new TransPath algorithm; a CNN-transformer-based hybrid with a token-aggregating and excitation module. SDSCL was effective with limited training samples, achieving higher results

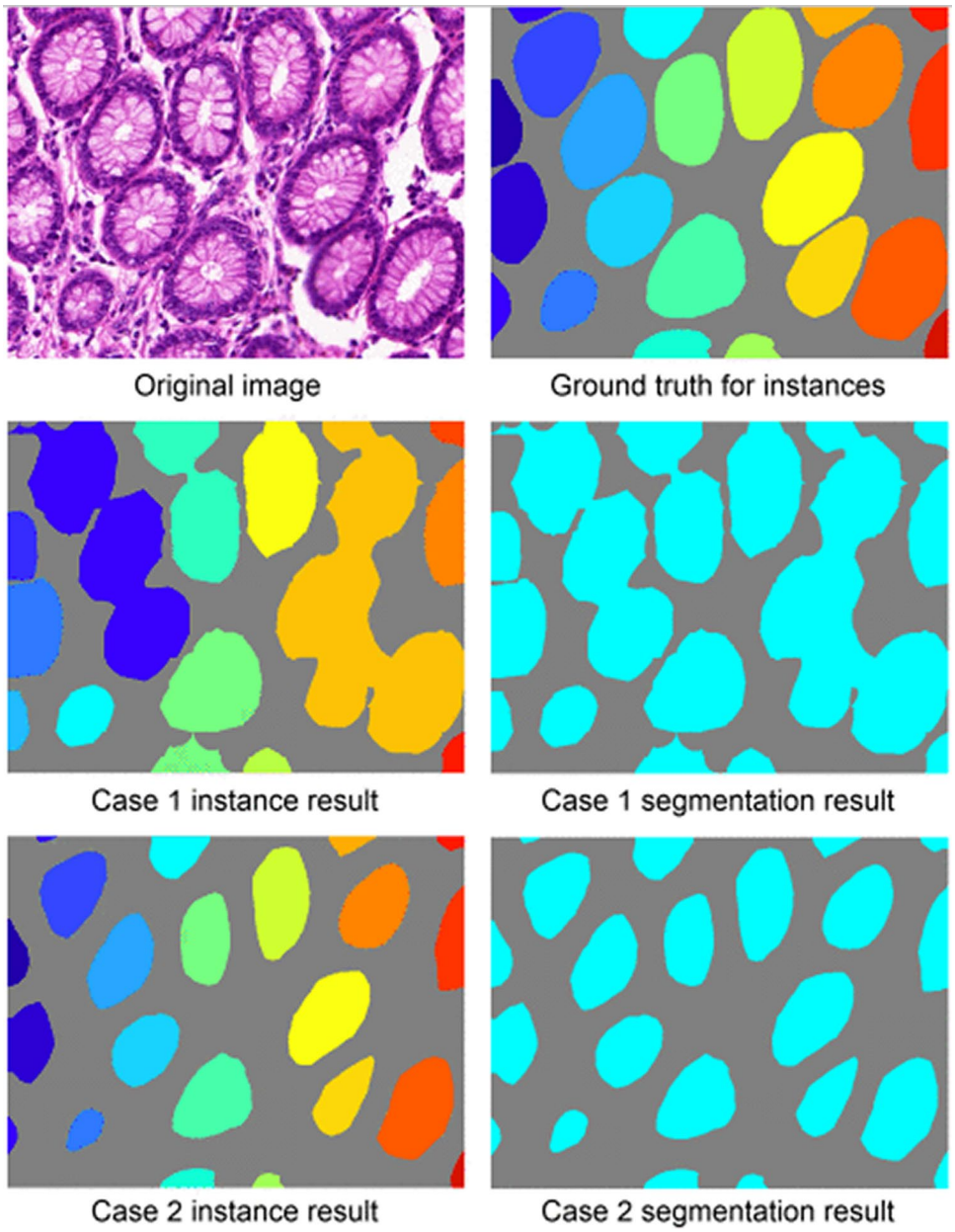

Fig. 3 Case 1 – glands are separated from the background, but many instances are not recognized. Case 2 – instances are labelled but many gland pixels are neglected [30]

using single classifiers, although slightly lower ones with multi-classifiers, offering a solution for scarce annotated data in clinical practice.

Abousamra et al. applied supervised DL to identify tumour-infiltrating lymphocytes (TIL) in gigapixel WSIs from TCGA [34]. Their dataset, spanning 23 cancers with 7983 WSIs, is the largest for TIL identification, addressing TIL heterogeneity across tissue samples. After selecting WSIs, labelling patches and annotating TIL-significant regions, they applied a DL model with three separate CNN architectures (VGG16, Inception-V4, and ResNet-34)to characterised samples as either Low, Medium, or High TIL based on the TIL-positive area ratio. The correlation between pathologist interpretation and the CNN score was observed. Pathologist annotations were combined with model predictions for cancers with scarce manual annotations, enhancing results. This approach offers utility in cancer immunotherapy, characterizing TIL abundance and spatial distribution to explore clinical significance and supporting pathology digitization in the future.

Sarker et al. used a CNN-based BC classification method using fused mobile-inverted bottleneck convolutions (FMB-Conv) with dual-squeeze and excitation (DSE) [35]. Their method of classification was robust: initially binary (benign or malignant) and then multi-class (adenosis, fibroadenoma, phyllodes tumour, tubular adenoma, carcinoma, lobular carcinoma, mucinous carcinoma, and papillary carcinoma) subtypes. Their model outperformed ResNet101,

InceptionResNetV2, and EfficientNetV2 networks on the BC BreakHis dataset. Attention mechanisms captured relevant features including cell structure, textural patterns, and morphology information, ignoring irrelevant background pixels. The model achieved high confidence rates but was limited by artefacts (blurriness, improper cell boundaries) in low-quality images.

CNN-based pixel prediction methods such as those by Ciresan et al. [13] and Wang et al. [17] utilised small image patches for mitosis detection but struggled with class imbalance, heterogeneous datasets, and staining artefacts. Sirinukunwattana et al. [18].improved nuclei detection by combining handcrafted features with CNNs while Xu et al. [30] used SSAE for nucleus detection, refining probability mapping techniques. SSAEs better-tackled issues around generalizability and enhanced detection sensitivity through the compression of high-dimensional histopathological image data into lower-dimensional features. They also learned hierarchical patterns from majority and minority classes making them better at handling class imbalance. However, CNNs with handcrafted features had enhanced performance through better incorporation of domain knowledge. Zhang et al. [26] integrated CNN-RNN models in urothelial cancer diagnosis to generate diagnostic descriptions while Graham et al. [27] enhanced instance segmentation with FCN-based methods for nuclei detection. Xu et al. [30] further improved gland segmentation using multi-channel DNNs but faced challenges in separating gland structures. Gong et al. [34] introduced SDSCL for CNN0based CAD for feature representation in data scarcity while Abousamra et al. [35] applied supervised DL to identify TILs in the largest dataset of its kind to support immunotherapy research.

Within the current landscape of supervised learning, key challenges persist due to class imbalances: rare events like mitosis detection suffer especially, leading to CNNs favouring non-mitotic classes. Varying staining and preparation protocols further affect generalizability which is compounded due to processing complaints that demand down-sampling and patch-based approaches. Future directions should emphasise WSI processing to help cancer grading, tumour localisation, and feature extraction at a large scale. Advancing architectures may also enable fine-grained subtype classification more accurately with the future of supervised learning in computational histopathology lying in hybrid approaches combining CNNs, transformers, contrastive learning, and domain-specific handcrafted features for maximized diagnostic accuracy.

2.2 Weakly supervised models

Hou et al. used multiple-instance Learning (MIL) to combine evidence from multiple patches sampled from the same image [36]. Their framework integrated MIL with CNNs, treating patches as instances with only image-level labels. The algorithm identified discriminative patches in high-resolution images, using Gaussian smoothing (edge detection and noise reduction) [37]. Applying this to WSIs from TCGA, they classified glioblastoma and LGGs with 97% accuracy, although subdividing LGGs demonstrated a much lower accuracy of 57.1%.

Jia et al. built an end-to-end learning system using a MIL framework connected with deep weak supervision (DWS) to segment cancerous regions in histopathology images [38]. They introduced constraints for positive instances to utilise additional weakly supervised information, enhancing the learning process. This method outperformed pathologists and could be applied across various histopathology imaging applications.

Liang et al. used weakly supervised DL for image segmentation by reiterative learning, proposing a neural network with overlapped region forecasts in gastric cancer to automate segmentation [39]. This method demonstrated superior performance without further manual annotation to train simple networks on weakly annotated biomedical images. Using two different patch extraction methods, predictions could be integrated, with weak annotations for improved data quality, achieving 91.09% mean accuracy, validating its cost-reduction potential. Figure 4 taken from Liang et al. [39] demonstrates their results but highlights the need for improved recognition of smaller cancerous regions and addressing *"over-fitting"* caused by excessive iterative learning.

Wang et al. used recalibrated multi-instance deep learning (RMDL) for whole-slide gastric image classification, utilising a two-stage framework, that accounted for varying instance contributions for final image-level label predictions [40]. RMDL captured instance-wise dependencies, recalibrating based on fused-features coefficients. Using a whole-slide gastric histopathology dataset with detailed pixel-level annotations, experimental results demonstrated RMDL outperformed other methods by > 4%, exceeding Liang et al. through considering each instance's impact [40]. However, large-scale annotation efforts and the need for compression methods to accelerate detection remain areas for improvement.

Challenges in histopathology arise through image-analysis variability. Standard instance-level aggregation methods, like voting or pooling, cannot guarantee accurate image-level predictions due to instance-level label misclassifications. Hou et al. used expectation–maximisation with CNN to output patch-level predictions trained via logistic regression for glioma subtype classification [41]. Campanella et al. used RNN to integrate semantically rich features across patch-level

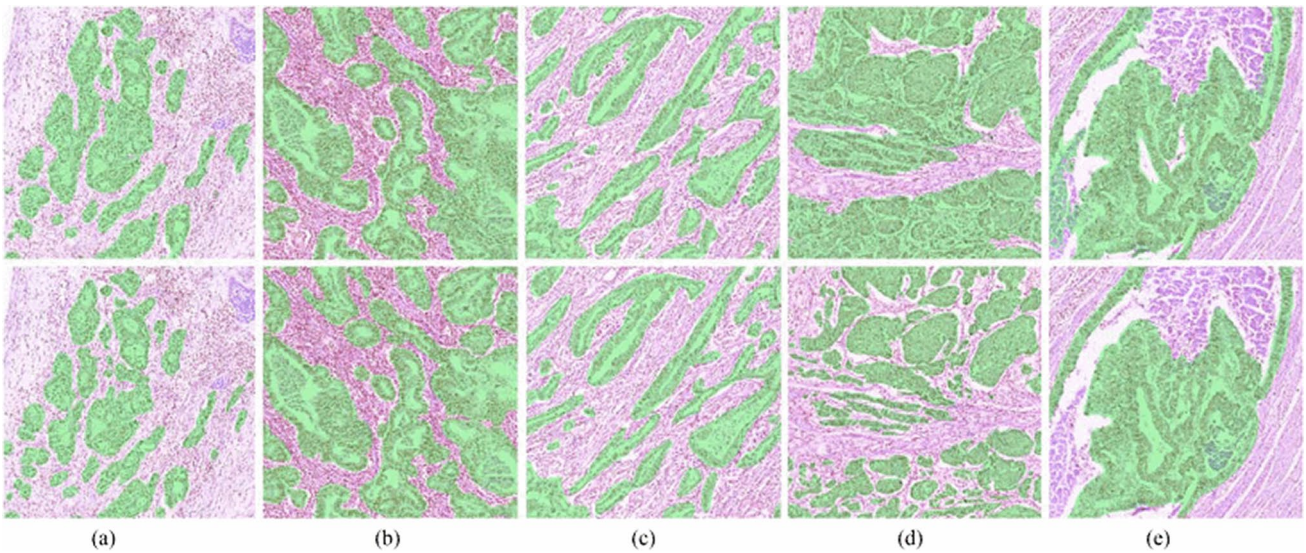

Fig. 4 Qualitative segmentation results on the dataset: top row – visualisation of the model, bottom row – visualisation of the ground truth. **a-c** well-segmented samples. **d** Gaps between cancerous areas with dense, small regions are ignored. **e** area of inflammation and necrosis are misdiagnosed in the lower region [39]

instances for final slide-level diagnosis [42]. The global–local detection method, popular in MIL, is clinically relevant for histopathological diagnosis by discriminating cancerous instances [43]. However, instance-level predictions often lacked validation, relying instead on heatmap visualisation due to annotation challenges.

MIL is one of many weakly-supervised learning methods. Others include models using minimal annotations at specific points [44, 45], bounding boxes [46], and partial pixel-level cancerous region annotations [47]. These methods propose variants to facilitate cost effectiveness and feature encoding strategies to derive labels from weaker annotations enabling fully supervised model training in a weakly-supervised manner.

Weakly-supervised learning is limited compared to fully-supervised learning. Guo et al. used a two-stage, weakly-supervised nuclei segmentation model requiring only nuclear centroid annotation demonstrating competitiveness. It combined segmentation networks, attention mechanisms, and low-level feature constraints outperforming other weakly-supervised methods, suggesting that further improvement is necessary for non-centroid points in histopathology segmentation [48].

Zheng et al. [49] applied a weakly supervised "human–machine fusion" model in renal cell carcinoma (RCC) histopathology, demonstrating superiority in clear-cell RCC classification. A Self-Supervised Clustering-constrained-Attention Multiple-instance (SSL-CLAM) model automatically detected diagnostically valuable subregions. Independent diagnoses by a junior pathologist (A) and expert uropathologist (B) on 445 WSIs repeated five times showed SSL-CLAM achieved an average accuracy of 0.787 (95% CI 0.772–0.801) (A) and 0.856 (B) (95% CI 0.843–0.867), highlighting human–machine fusion's advantage. Further prospective trials are necessary, however, to validate this strategy's effectiveness. SSL-CLAM is not yet suitable for independent diagnostics, but has an adjunct utility: requiring standardised slide preparation for better output.

MIL-based approaches (Hou [36], Jia [38], Wang [40]) showed efficacy for histopathological classification and segmentation but required enhancements for instance-level variability and small region recognition, often requiring additional supervision constraints or being prone to overfitting. Comparatively, weakly-supervised models that applied reiterative learning [39] also faced challenges like overfitting and dependency on structured annotations. While Human–Machine Fusion [49] has the potential to aid pathologist decision-making, standardised protocols are necessary before independent application. Importantly, weakly supervised learning methods offer cost-effective alternatives to fully supervised models, although they currently require improvements in validation and efficiency. Future work could emphasise data augmentation and knowledge distillation approaches to balance annotation costs while enhancing human–machine fusion through interactive AI systems to enable more reliable collaboration. However, the increasing use of AI would require prospective clinical trials and governance approval with standardised processing pipelines and structured compliance guidance with benchmark metrics for essential real-world use.

2.3 Unsupervised models

Xu et al. employed unsupervised DL using an SSAE for nuclei detection on high-resolution histopathological images of BC [20]. This method learned high-level features from pixel intensities alone to identify distinct nuclei features. A sliding window operation represented image patches obtained via the auto-encoder and was fed into a classifier categorising each patch as nuclear or non-nuclear. The SSAE achieved an improved F-measure of 84.49% and an average AUC-PR(Area Under Precision-Recall curve) of 78.83%, outperforming several other nuclear-detection strategies. Focussing specifically on extracting nuclear boundaries, this approach could enhance segmentation models such as watershed [50], active contour [51], and region-growing approaches [52]. Xu et al.'s framework enabled cell-by-cell graph feature development to assess cellular topology in tumour histology [20].

Hu et al. applied label-free unsupervised classification with interpretable visualisation for bone marrow (BM) cellular components [53]. They exploited cell elements using BM datasets combining nuclei segmentation with cell-level visual representation to highlight cellular varieties (Fig. 5) [53]. Using a unified Generative Adversarial Network (GAN) with a formulation of loss of function inherits both WGAN-GP (Wasserstein Generative Adversarial Network with Gradient Penalty) and InfoGAN (Information Maximising Generative Adversarial Network). They performed unsupervised classification and interpretable visualisation by maximising mutual information. However, improvements of the segmentation method and computational efficiency are needed. Additional information, including clinical trial and gene expression information, is required to re-evaluate risks after annotating relevant genetic variants.

Bulten et al. [54] used unsupervised learning to create a self-clustering convolutional adversarial autoencoder for prostate cancer (PC) classification without prior training. The clustering adversarial autoencoder (CAAE) clustered tissue during training, eliminating the need for post-processing methods like K-means or t-SNE. Trained on patches extracted from 94 registered WSI pairs (54 training, 40 test)post radical prostatectomy, PC was identified by the presence of epithelial markers and absence of basal cell markers. The CAAE achieved an F1 score of 0.62 when discriminating tumour vs non-tumour outperforming the control (H&E to H&E; 0.52) demonstrating that the reconstruction task had benefit in learning cross-domain mapping. A flaw was the dependence on sample quality: with noisy, heterogeneous data limiting results. Future work should expand the autoencoder's field to visualise whole images instead of patches.

Sari et al. presented an unsupervised learning model to classify histopathological tissue by quantizing salient subregions via feature extraction from image data [55]. Restricted Boltzmann machines (RBMs) defined hidden unit node activation values in the final RBM as features, clustering colon tissue images for representation and classification. The LearnDBN primary model was used to learn the deep belief network of RBMs and define the local deep feature. The secondary model, LearnClusteringVectors, clustered vectors onto local deep features. H&E-stained datasets with 1644 images (510 normal, 859 low-grade (LGC) cancerous and 275 high-grade cancerous) with a further subset of the first with low-grade cancerous was sub-categorized to solve difficult sub-categorization; labelling them into five classes: normal, LGC 1, LGC 1/2, LGC 2, and high grade. The results were promising, although challenges included potential misclassification for large, non-epithelial regions and requiring low magnification when utilised as a whole-slide scanner.

Recently, Sheikh et al. applied unsupervised learning for WSI diagnosis with stacked autoencoders processing multiple image descriptors (histograms of oriented gradients, local binary patterns, and fused heterogeneous features) [56]. Latent

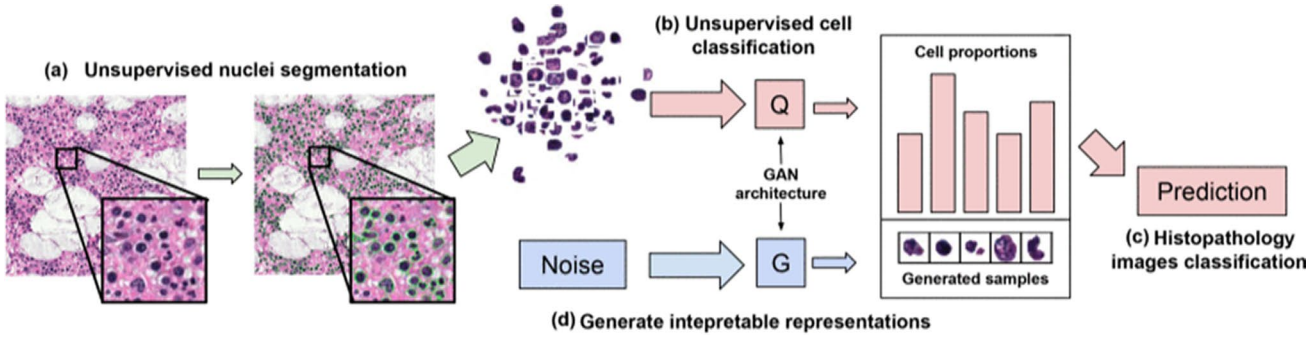

Fig. 5 Overview of the pipeline: **A** nuclei segmentation performed on histopathology images. **B** trained GAN architecture, clustering performed using a learned auxiliary network. **C** Image-level prediction based on cell proportions. **D** generator G can generate an interpretable representation for each category [53]

vectors were extracted from individual autoencoders and fused feature representations were used for classification. Training with additional descriptors helped identification of intricate cellular structure, out-performing state-of-the-art approaches, achieving the highest accuracies of 87.2 and 94.6 for ICIAR2018 and Dartmouth datasets respectively. The model also classified breast and lung cancer types comparable to pathologist visualisation. However, the model was limited in determining optimal patch size for robust representations in the bottleneck space and separating morphologically heterogeneous regions in sub-multi-level tumours due to unbalanced data (31 WSIs), affecting the feature fusion stage. Future works should utilise larger, balanced patch datasets to improve robustness.

Nam et al. evaluated a pre-operative CT-based DL model using unsupervised learning cluster analyses to identify histopathologic risk factors in 1667 patients with resected Stage I-IV lung adenocarcinoma [57] in a retrospective study. The model showed associations with aggressive adenocarcinoma subtype (odds ratio (OR), 1.03; $P = 0.03$), venous invasion (OR, 1.0, $P = 0.02$), and visceral pleural invasion (OR, 1.08, $P < 0.001$) when adjusted for clinical T category and radiologic nodule type. The model also predicted patient survival and identified associations with histopathologic and metabolic findings from preoperative chest CT scans demonstrating utility when identifying histopathologic findings indirectly and for survival prediction. Key limitations included single-centre data, retrospective design, and sole analysis of epidermal growth factor receptor (EGFR) mutations. However, it highlighted the possibility of using unsupervised learning to enable post-operative survival prediction but required multi-centre studies for validation.

Niehues *et* used attention-based self-supervised learning in colorectal cancer (CRC) demonstrating success in identifying biomarkers from WSI (including microsatellite instability (MSI), BRAF, KRAS, NRAS, and PIK3CA) outperforming previous approaches [58]. An attention-based approach combines global features and minimises artefacts; demonstrating efficient DL prediction and location of biomarkers in WSIs cost-effectively. However, its success is limited to identifying MSIs; with KRAS, NRAS, and PIK3CA not achieving clinically relevant ranges. Nonetheless, it demonstrates independent transformation of unsupervised learning into supervised learning through the auto-generation of labels.

Hosseini et al. addressed real-world data imbalances by modifying existing DL models [59]. Their approach transferred knowledge from label-rich sources to unannotated targets using focal loss functions to tackle class imbalance challenges and co-training with pseudo-labelled target data. Applying this to PC images, they classified them into low-cancerous and high-cancerous regions, assigning labels to unlabelled data, selecting high-compatibility targets for training and demonstrating efficacy on public datasets.

Fetisov et al. propose a novel unsupervised meta-learning-based segmentation method which corresponded to classification tasks and eliminated the need for patch-level annotation [60]. It determined the relative importance of tissue regions while utilising the PANDA dataset for unsupervised segmentation, capable of detecting malignant regions with a 0.79 patch-based AUC (area under the curve). This fully unsupervised model aids machine learning approaches in cancer histopathology utilising segmentation as a feedback mechanism to identify regions needing scrutiny. The precision of segmentation, however, was limited by patch size.

Wang et al. [61] introduced a Transformer-based SSL strategy, Semantically-Relevant Contrastive Learning (SRCL) with an integrated hybrid CNN-Transformer (CTransPath) to improve feature extraction in histopathology. A 15 million patch, 30,000 WSI dataset was used for pretraining and the model was validated on five tasks: patch retrieval, classification, weakly-supervised WSI image classification, mitosis detection, and colorectal adenocarcinoma gland segmentation. Combining CNN's local texture extraction and the global feature-learning ability of Transformers improved classification and segmentation performance. This also allowed further generalisation of downstream tasks allowing universal feature extraction for histopathology. However, it came with a high computational cost burden, making deployment challenging in low-resource settings. There was also potential for incorrect positive pairing: while SRCL selects semantically similar patches, false positives may still occur that could reduce the effectiveness of contrastive learning.

Dippel et al. [62] introduced RudolfV for digital pathology which integrated pathologist knowledge into the development process. Using a diverse dataset of 103,849 WSIs from various labs, with various staining protocols (H&E and non-H&E) the authors curated a semi-automated process using the DINOv2 framework and evaluated RudolfV on multiple public and internal benchmarks to demonstrate competitiveness. RudolfV has applications in digital pathology (cancer detection, tissue classification, and biomarker scoring) and biomedical research to serve as a foundation for the development of diagnostic tools and insights into disease mechanisms. However, computational complexity and significant overlap through the use of multiple datasets uses significant resources which limit accessibility for smaller labs and risk overlapping data with the pre-training dataset that may potentially inflate performance metrics. Future outlooks could focus on scaling the dataset to handle more complex tasks and integrate the model into clinical workflows to assist pathologists in real-world diagnostic settings. RudolfV may benefit from extending the model to address rare diseases with limited utility to further enhance its clinical practice and integration into clinical workflows.

From these selected examples, we can ascertain that unsupervised learning has significant capacity in computational histopathology aiming to reduce manual annotation dependency and enhance feature extraction. Recent advances including contrastive learning, GANs, and hybrid CNN-Transformer architectures have led to advances in nuclei detection, cancer classification, biomarker identification, and segmentation. Unsupervised learning possesses adaptability for multiple pathology tasks, potential for survival prediction and clinical integration, and enhanced segmentation approaches which improve methods like watershed, active contour, and region-growing. However, unsupervised learning currently has limited interpretability (thus requiring clinical validation) and higher false-positive rates where methods like SRCL and contrastive learning introduce incorrect positive pairing which reduces classification accuracy. Finally, the patch-based approach loses broader spatial relationships which are identified through WSI analysis.

Future research should emphasise WSI analysis with multimodal AI integration to combine pathology images with genomics, radiology, and clinical metadata while addressing class imbalance issues with balanced datasets for improved generalisation. Unsupervised learning has the potential to play an important role in AI-assisted pathology workflows supporting diagnostics, research, and clinical decision-making but is currently limited due to expense and limited reliability when resources are scarce.

Transfer learning remains the predominant method in digital pathology, leveraging knowledge from one domain to another (source to target), to address limited target training data [63]. It emphasises transferring knowledge with or without labelled data, spanning all levels of supervision. In histopathology, ImageNet pre-trained models such as VGG-Net [64], InceptionNet [65], ResNet [66], MobileNet [67], DenseNet [68], and other variants are commonly used (Table 5).

Tsuneki et al. [69] used transfer learning to classify prostate adenocarcinoma needle biopsy WSIs into adenocarcinomatous and benign lesions. Among three trained models, TL-colon poorly ADC-2 (20x, 512) demonstrated superior performance with ROC-AUCs of 0.967–0.978, 0.737–0.910 and 0.987 on TUR-P and TCGA datasets. Automation could reduce the interpretative burden on pathologists, minimise missed cancer foci, and systematically handle substantial WSI volumes without fatigue-associated bias. Further validation could potentially transform precision oncology.

Noorbakhsh et al. [70] used transfer learning to quantify tumour cellularity (TC) to assess tumour burden (TB) in BC measuring response following neoadjuvant therapy. Transfer learning outperformed pathologist review (93% to 81%) surpassing hand-engineered approaches prone to misclassification of malignant nuclei. Three TC scoring methods were tested: visual interpretation of TB, a pathologist-mimicking approach (identifying malignant: benign cells and stroma), analysing cell proportions, and automated feature extraction using deep CNN. limited due to being time-consuming and human fatigue. Automated techniques produced TC scores similar to trained pathologist reviews and performed better when identifying healthy tissue and cancerous patches, addressing the limitations of traditional manual evaluation.

Due to the vast number of features used for decision-making, class-specific discriminative regions via activation maps identify regions responsible for presentations. Mahmood et al. [71] used Faster-Regional CNN (FR-CNN) to create a multi-stage mitotic-cell-detection method using ICPR (International Conference on Pattern Recognition)-2012 and ICPR-2014. Activation maps and routing through different layers of trained ResNet-50 and DenseNet-201 classifiers allowed multi-dimensional layers to present as a single image. Layering allows a combination of simple (colour, edges etc.) and deeper complex features and significantly impacts decision making. Figure 6 demonstrates the characteristic activation maps of mitotic vs non-mitotic [71].

Advances in DL decrease the gap between pathologists and computer performance. Significant variation in mitotic cell size highlights FR-CNN for efficient feature extraction. Mitotic scores, reflecting tumour proliferative activity, are obtained by counting mitotic figures in hotspots [72–74]. DL, however, is highly reliant on large datasets necessitating data augmentation with smaller sample sizes. posing challenges distinguishing variation between mitotic and non-mitotic cells. Currently, pathologists outperform AI in decision-making suggesting FR-CNN is best suited as an augmentative tool.

Pantanowitz et al. [75] trained an AI model to detect mitoses using 320 breast invasive ductal carcinoma images. 24 readers were assessed 140 high-powered fields with and without AI comparing accuracy and efficiency. AI intervention improved precision and sensitivity, with 87.5% identifying more mitoses and 54.2% demonstrating lower rates of false positives. AI completed tasks 27.8% faster, enhancing accuracy and efficiency in invasive breast carcinoma quantification.

Kather et al. [76] used DL to infer genetic mutations, molecular tumour subtypes, gene expression signatures, and biomarkers from routine histology in across > 5000 patients. 14 common solid tumour types were tested: breast (BRCA), cervical (CESC), colorectal (COAD and READ), gastric (STAD), head and neck (HNSC), hepatocellular (LIHC), lung adenocarcinoma (LUAD), lung squamous (LUSC),melanoma (SKCM), pancreatic (PAAD), prostate (PRAD), renal cell chromophobe (KICH), renal cell clear cell (KIRC), and renal cell papillary cancer (KIRP). Their workflow retrained a deep neural network for each target using identical parameters. Following WSI scanning and patient-level cross-validation (splitting the cohort into three partitions, neural networks trained on two partitions, and evaluated on the third), deep neural networks

Table 5 Key examples of transfer learning methods in histopathology summarised (ICPR = International Conference of Pattern Recognition)

Study	Cancer types	Stain	Application	Method	Dataset
Tsuneki et al. [69]	Prostate adenocarcinoma	H&E	Prostate adenocarcinoma classification in needle biopsy WSIs	TensorFlow CNN with transfer learning and bootstrap method	TCGA
Noorbaksh et al. [70]	Multi-cancer	H&E	Pan-cancer classification	Pre-trained inception-V3	TCGA (27,815 WSIs)
Mahmood et al. [71]	Breast	H&E	Detection of mitotic cell count	Faster Regional CNN (FR-CNN)	ICPR2012 ICPR2014
Pantanowitz et al. [75]	Breast invasive ductal carcinoma	H&E	Mitotic cell count	FR-CNN by ResNet-101	320 WSIs from UPMC and SMC
Kather et al. [76]	Multi-cancer	H&E	Molecular alteration detection using MSI detection as a benchmark task	Hyperparameter selection by ResNet-18, AlexNet, Inception-V3, DenseNet-201 and ShuffleNet	> 5000 WSIs from multiple institutions
Figueira et al. [77]	Lung, breast, colon cancer	–	Identification of lung, breast, colon cancer using cross-domain adaptation	Fusion of ABDA-Net with ResNet for inter-dataset transfer of knowledge	Camelyon 16 (breast) – 649,045 samples, ACDC LungHP (lung) 200,000 samples, DigestPath (colon) 147,210 samples

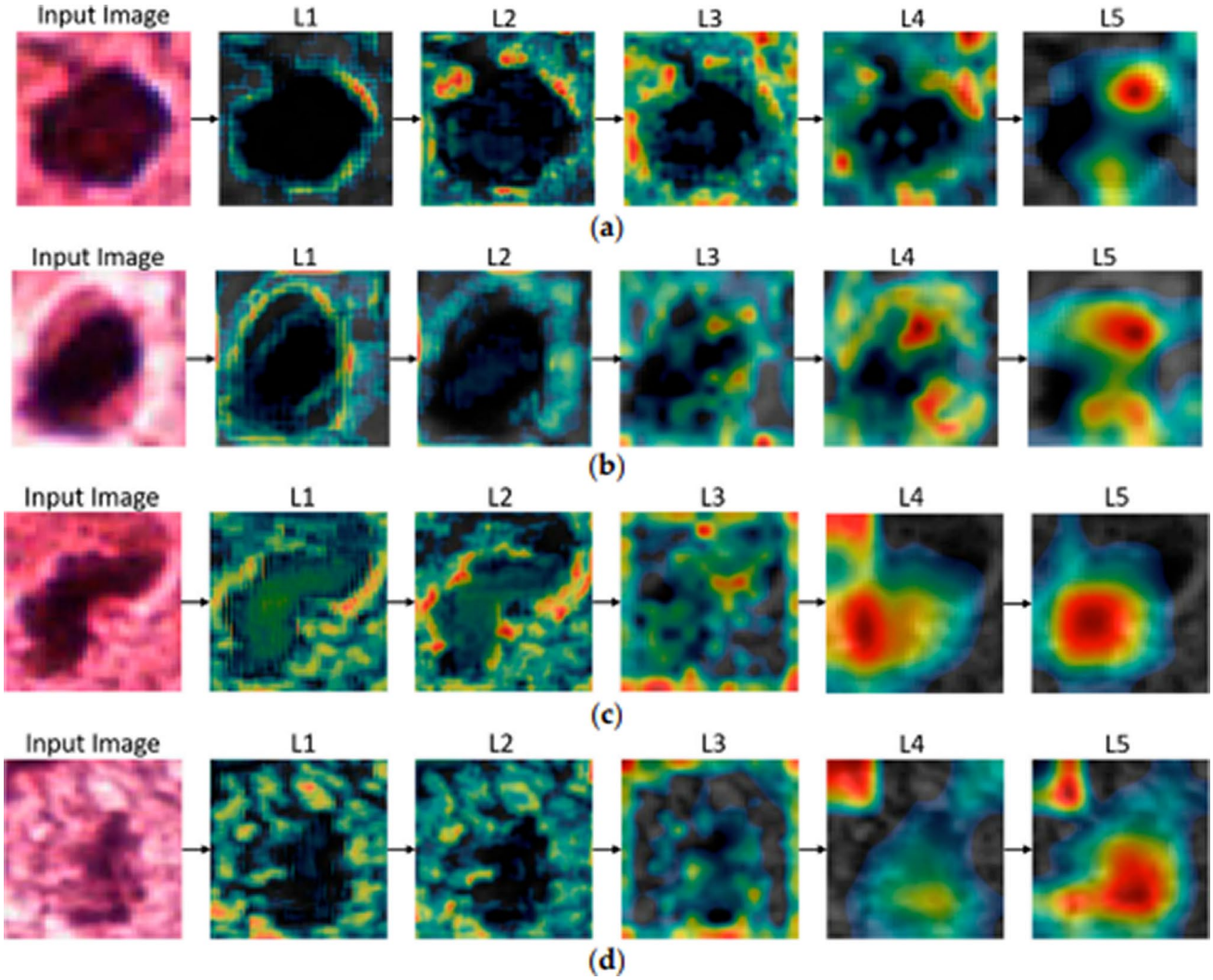

Fig. 6 Activation maps from different parts of (**a**), (**b**) ResNet-50 and (**c**), (**d**) DenseNet-201 with mitotic and non-mitotic cell images. Mitotic cells were demonstrated at (**a**) and (**c**) and non-mitotic cells shown at (**b**) and (**d**) [71]

predicted molecular status for each tile with the percentage of positive-predicted tiles for each class displayed as a probability score for each patient. Their method detected subtle patterns aiding classification, identifying basal BRCA subtypes, high/low proliferation LUADs, and STAD but was limited due to reduced data availability for a multi-cancer approach.

Despite being widely utilised in digital pathology, the effectiveness of transfer learning can vary and is dependent on numerous factors including task complexity, dataset size, and computational constraints.

High computational cost especially in real-time clinical applications remains a significant burden. Figueira et al. [77] proposed a method for lung, breast, and colon cancer detection using domain adaptation networks (ABDA-Net) combined with ResNet50 as a feature encoder with their method achieving an 86.61% accuracy on the lung cancer dataset. Their cross-domain adaptation technique uses adversarial learning to transfer knowledge between datasets integrated supervised learning (for domain training) and unsupervised learning (for domain adaptation) in a multi-modal approach. Further work could integrate multiple data types beyond histopathology while also integrating clinical and patient-specific data. Fusion with genomics (multi-omics integration) could enable molecular subtyping of tumours and improve tumour classification and prognosis prediction by linking tissue morphology with underlying genetic mechanisms using neural networks to integrate structured genomic representation. Fusion with radiology (radio-pathomic) could integrate imaging with DL fusion models to link imaging phenotypes, molecular markers, and histopathology data to improve tumour staging and therapy planning. While Figueira et al. [79] proposed a method for cross-domain adaptation of

histopathological images only, true multimodal extension would incorporate clinical, genomic, and radiologic data to enable personalised tumour diagnosis and treatment prediction making the AI system more clinically relevant.

3 The future of AI in computational histopathology

Overall, AI application to histological image analysis has gained momentum, with AI application for WSI normalisation, identification of ROIs and overall classification of images into specific classes. Combining digital histopathology and AI presents an opportunity, in the current paradigm, as an aid to pathologists' tasks with a final long-term goal of minimising pathologist input maximally.

3.1 Data fusion

Clinical and histopathological characteristics alone cannot reliably predict carcinoma progression. Integration of multi-source data is vital for tailored diagnosis, with molecular features gaining increasing value in identifying cancer type, progression, demographics, and prognosis. AI algorithms enable extensive information extraction from WSIs including proteomics, immunohistochemistry (IHC), genetic and epigenetic data, spectroscopic images, and large-scale imaging (MRI, CT, etc.). This could aid in predicting features of poorer prognosis and progression in malignant cancer e.g. through integrating current knowledge about molecular mechanisms in cancer metastasis. Furthermore, -omics research can be combined with spectroscopy for identification and localization of pathophysiological cell and tissue alterations. Path-omic fusion [78] offers end-to-end multimodal fusion of histology image and -omics (mutations, copy number variations (CNV), RNA-Seq etc.) data to predict survival. Combining deep learning with datasets from available databases; including information such as WSIs, genotypic, and transcriptomic data; using multimodal fusion (MMF) to allow faster prediction on an efficient individual-to-individual basis.

Chen et al. [78] developed a multimodal data fusion method for pan-cancer integrative histology-genomic analysis using DL; demonstrating improved prognostic models for most cancer types, jointly examining WSIs and molecular profiles from 14 cancer types. Using a weakly-supervised DL algorithm, Chen et al. fused heterogenous modalities to identify prognostic features. MMF algorithms combine H&E WSIs and molecular profile features (mutation status, copy-number variation, RNA-Seq expression etc.) to measure and explain relative risk of cancer death. They demonstrated that MMF applied to AI tasks (including survival-outcome prediction, histopathology features, and molecular features) using weakly-supervised learning allowed predicting interactions and identifying low-risk/high-risk patients. The study used a 6,592 giga-pixel WSI dataset from 5,720 samples across 14 cancer types from TCGA with specific training for each type. This allowed investigation of local and global image-comics and quantification of tissue microarchitecture to identify shifts in feature importance between unimodal and multimodal [78]. Figure 7 demonstrates this [78].

Chen et al. [78] demonstrated proof-of-concept from multiple data sources to identify correlative features driving prognosis, enabling future work using larger multi-modal cohorts. Robust characterization and spatial organisation to address micro-problems including intra-tumoral heterogeneity, immune cell presence, and morphological variation is necessitated to maximise opportunities for data fusion and integration.

3.2 Swarm learning

Access to relevant data is a major drawback for many medical applications due to strict privacy regulations. Full anonymization of data is challenging, but connecting follow-up information with a specific data point is possible. Smart contracts and blockchain technology could securely share data based on authentication, allowing expert opinions or running samples on other diagnostic systems. This approach benefits computational pathology through secure, efficient role-based data sharing.

Saldanha et al. [79] attempted to integrate blockchain technology into imaging-based cancer molecular diagnosis tackling the limitation of large data size requirements and barriers to data via a decentralised multicentric model training: Swarm Learning (SL). This pilot highlighted the predictive accuracy of BRAF mutational status and microsatellite instability (MSI). Using small datasets in medical AI is challenging as prediction performance increases with larger training datasets. Saldanha et al. [79] demonstrated that SL enables AI-based prediction of BRAF and MSI status in CRC and for other image classification tasks in computational pathology.

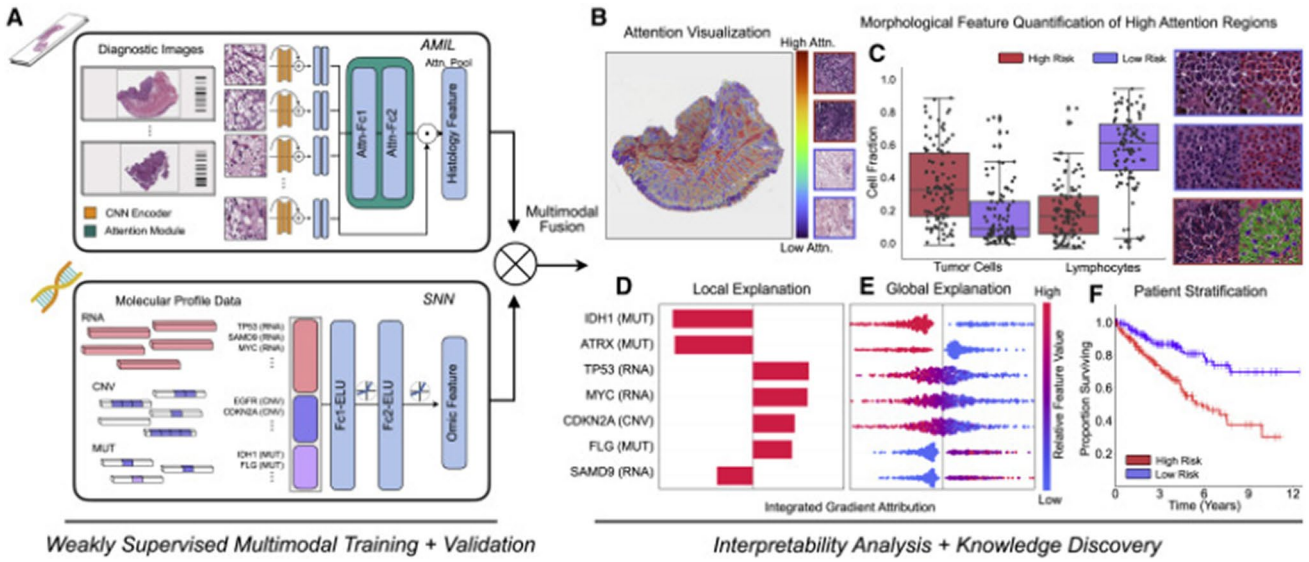

Fig. 7 **A** Patient data in digitised high-resolution H&E histology. **B** Explanations per-patient as heatmaps. Red = high attention regions (contributing to predicted risk score). **C** Global morphological patterns extracted via cell quantification of high-attention regions in low and high-risk patient cohorts. **D** Explanations per-patient using attribution-based interpretability in integrated gradients. **E** Global interpretability for molecular features showing feature value and magnitude of gene attributions across all patients. **F** Kaplan–Meier analysis for patient stratification [78]

Training AI models with small datasets can be as effective as using a large dataset; potentially lowering hardware requirements, reducing cost burden which is beneficial in low/middle-income countries. However, the study concluded that their model did not improve precision and accuracy for the two tasks specifically: dependency on local models was suspected to be the cause of their advantages necessitating future studies, necessitating future studies to ensure tangible benefits of blockchain and AI integration The study may be limited as differential privacy was not explicitly investigated but allows potential for future incorporation. Swarm learning provides a key step forward in medical AI by allowing multiple devices to implement pattern recognition code with open-source software, feasible across multiple healthcare institutions. Saldanha et al. [79] provide clear guidelines to establish SL in routine workflows.

SL may face challenges related to data heterogeneity, computational demands, and model generalization [80]. Differing institutional protocols may lead to domain shifts across datasets. SL instead relies on distributed, locally trained models, making aggregation to a generalizable global model difficult. This also raises a risk of bias with non-identically distributed data. Institutions with differing patient demographics, limited data size, limited cancer subtypes, or sample distributions can bias the model leading to poor convergence and inconsistent performance when applied externally. This may also lead to overfitting on institutional-specific patterns where the model learns site-specific artefacts rather than general pathology features, therefore failing on external datasets. There is scope for hybrid approaches combining SL with self-supervised learning and domain adaptation to enhance model robustness and clinical applicability.

3.3 Interpretability

Clinicians often resist adopting new AI technology due to lack of understanding, even with evident benefits. "Black-box" essentially describes the challenge of explaining AI models. Bhattacharjee et al. explored explainable computer vision (XCV) using a DL model to classify, differentiate, and PC [81]. To demonstrate transparency of their model decision-making they presented three techniques: Activation Layer Visualization (ALV), Local Interpretable Model-Agnostic Explanation (LIME), Shapley Additive exPlanations (SHAPE and Gradient-weighted Class Activation Mapping (Grad-CAM). XCV is described as "AI in which the results of the black-box model can be understood by humans". Their Light-Dense CNN (LDCNN), modified from light-weight CNN (LWCNN) classifies H&E-stained images of PC showing superiority for tissue image classification and cancer detection. Their work warrants further research applied to other cancers and real-life health scenarios.

3.4 three-dimensional (3D) histopathology

Technological advances now allow imaging of large tissue samples in 3-D mimicking H&E slides. This approach can detect diagnostically significant tissue that's difficult to see in 2-D sections, such as lympho-vascular invasion, minimal residue, and tertiary lymphoids [82]. In PC, 3-D imaging better characterises the heterogeneous and irregular nature of the tumour microenvironment (TME); aiding therapy through a superior understanding of immune cell migration [83]. However, the complexity and large data volume make 3-D histopathology interpretation challenging. A recent study trained AI to handle 3-D image sections slice-by-slice, outperforming current 2-D histopathology AI applications [84].

3.5 Vision transformers (ViTs)

ViTs provide a DL designed for image classification tasks, using a transformer architecture that divides images into patches, processes through the transformer layers, and captures global relationships for effective pattern recognition. Chen et al. combined CNN and ViT using a parallel dual-branch network (DBNet) based on ResNet to capture local and global elements in histological images; showing improvement over state-of-the-art CNN models. through a multi-scale ViT approach for gastric cancer identification in histological images to produce diagnostic predictions by combining global and local features [85, 86].

Vorontsov et al. [87] implemented Virchow; a ViT model with 632 million parameters trained on 1.5 million H&E WSIs, significantly larger than prior models, processing massive datasets to create generalizable embeddings for diverse downstream tasks in pathology. Virchow demonstrated an AUC of 0.949 in pan-cancer detection across 17 common cancer types and 0.937 for rare cancers outperforming current state-of-the-art models. For biomarker detection, the Virchow model demonstrated consistently high performance for MSI in colon cancer (AUC 0.972; 95% CI 0.950–0.989), fibroblast growth-factor receptor (FGFR) in bladder cancer (AUC 0.902; 95% CI 0.862–0.941) and for EGFR in lung cancer (AUC 0.853; 95% CI 0.804–0.891). Vorontsov et al.'s model enables cancer, detection, subtyping, and biomarker prediction simultaneously while demonstrating adaptability for common and rare cancers. Their multi-view self-supervised learning model captured a broad range of pathological features including tissue morphology, nuclear, atypia, and inflammatory responses. However, their dataset was sourced from a single institution (MSKCC) potentially limiting generalisability from other institutions. They also acknowledged the need for aggregator architecture and training procedures given that their embeddings were generated at the tile-level. Nevertheless, they demonstrated that pan-cancer prediction remained robust from external sites, on out-of-distribution data, and on rare cancer types.

Chen et al. [88] introduced a general-purpose foundation model trained on 100 million images from over 100,000 diagnostic H&E-stained WSIs across 20 major tissue types. Their model (UNI) outperformed previous state-of-the-art models including CTransPath and REMEDIS across 34 representative tasks of varying diagnostic difficulty while also introducing novel capacities including resolution-agnostic classification and few-shot learning with superior generalizability. UNI introduced a prompting-based classification approach (MI-SimpleShot) which enhanced slide-level disease identification, increasing data efficiency through high accuracy achieved with reduced labelling. When compared to the 4-shot performance of UNI with other models, other models needed up to eight-fold training examples per class to achieve the same output [86]. The model possesses limitations in segmentation tasks and multimodal integration – key areas for further improvement. Their Vision Transformer Large (ViT-L) is not inherently optimised for dense prediction tasks such as segmentation: unlike CNNs, ViTs lack inductive biases for spatial structure, instead processing images as patches which limits their ability to capture fine details. Combining ViT with CNN-like hierarchical structures could offer a solution for improved segmentation in a hybrid model.

Lu et al. developed a multi-modal generative AI copilot for human pathology that combined a ViT (trained on 100 million images) with a large language model and then fine-tuned on 456,000 + visual-language pathology instructions which covered 999,202 question–answer pairs [89]. They evaluated their model (PathChat) against GPT-4 V, LLaVA-Med (a biomedicine-specific multi-modal large language model (MLLM)), and LLaVA 1.5 (a general purpose MLLM) to cover 54 diagnoses in 11 major pathology fields. Notably, their model had a 7.1% accuracy without context which improved to 89.5% with clinical data outperforming GPT-4 V significantly (26.9% without context and 53.8% with it). Their model had utility in pathology question benchmarking, generating more accurate and preferred responses

compared to other AI models as ranked by seven expert pathologists especially in microscopy analysis. Path Chat also has utility in education and clinical assistance, serving as a real-time pathology tutor and decision-support tool However, it had limited clinical and ancillary knowledge, being outperformed by GPT-4 V in general medical knowledge and biomarker analysis while also requiring an expensive deployment model with large computational resources. Currently, it cannot process WSIs instead relying on selective regions of interest (ROIs). Nevertheless, it represents a step in integrating AI copilots into pathology workflows with a future outlook to becoming an essential assistant in pathology, possessing utility in resource-limited settings and for pathologists-in-training.

3.6 Bias mitigation

Vaidya et al. [90] addressed demographic bias causing misdiagnosis in computational pathology models and investigated the performance disparities of DL across different demographic groups highlighting that publicly available datasets (e.g. TCGA) exhibited significant gaps with underrepresented demographic groups in key tasks such as cancer subtyping and mutation prediction. They notably identified performance gaps in cancer subtyping between white and black patients. The difference in AUROC was 3.0% for breast cancer, 10.9% for lung cancer, and 16.05 when identifying IDH1 mutations for gliomas.

Vaidya et al. [90] further concluded that unsupervised vision foundation models trained on large datasets reduced performance disparities between groups but did not eliminate gaps, necessitating further bias mitigation efforts. Site stratification would hold out certain sites during training to prevent models from incorporating certain site-specific biases which may show mixed results. Stain normalisation to reduce variability was also suggested – however, it only improved performance selectively. Test Set Resampling was used to ensure equal representation of demographic groups – nevertheless, performance disparities persisted suggesting complex causes of bias. Analysis beyond race (including income, insurance status, and age) still showed persistent disparities further highlighting the need for more comprehensive mitigation strategies. This study highlighted the necessity for enhanced regulatory oversight to require demographic-stratified performance for AI models in healthcare to ensure equitable outcomes. It also underscored the need for further research to understand the complex interplay between demographic factors especially in intersectional contexts.

Distribution shifts occur when data distribution changes, thus affecting ML models. Kulinski et al. [91] proposed a novel, practical framework to explain distribution shifts using interpretable transport maps to offer a balance between detail and interpretability. They provided actionable insights to aid human operators in making informed decisions. Computational histopathology presents distribution shifts through numerous sources of variation: staining protocols, scanner differences, or tissue preparation methods across hospitals and labs. These can impact the performance of ML models for cancer detection and tissue classification. The framework provided by Kulinski et al. [91] uses k-cluster transport to reveal how different subgroups of histopathology can be affected by distribution shifts which allow targeted data augmentation and domain adaptation techniques. This provides more interpretable mapping to produce actionable insights which can help improve the robustness and generalizability of machine learning models. Challenges persist, however, especially when handling complex content-based shifts and maintaining the interpretability of explanations. Further research should emphasise more domain-specific interpretable methods and causal explanations to utility in histopathology.

Demographic-dependent performance disparities and distribution shifts from varying protocols remain significant sources of bias which are especially problematic in sensitive fields like computational histopathology. Although significant strides are made to mitigate these, future research is needed to understand the complex interplay between demographic factors that produce variation and to develop fairer models that are less sensitive to biased features while providing actionable insights for reducing bias.

4 Conclusion

In summary, AI has made significant strides in cancer research, particularly in diagnosis and prognosis, with multiple models performing complex functions with increasing accuracy. Despite broad applications, considerable potential for further advancement remains. Challenges persist in transitioning AI technologies from research to clinical settings, with significant privacy concerns. This review offers a comprehensive overview of various AI models in computational histopathology from 35 studies. Supervised, weakly -supervised, unsupervised, and transfer learning approaches have demonstrated depth and proficiency. Through critical analysis of their strengths and limitations, we provide insights into

future perspectives and emerging technologies. Moving forward, we aim to explore additional subcategories of learning methods and offer comparisons between deep and shallow learning techniques.

Author contributions Conceptualisation and design—AT, AG, SA Data collection and assembly—AT, SSF Data analysis and interpretation—AT, SM, SB Manuscript writing—AT, AG, MH Final Approval of Manuscript—All Authors.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J Clin. 2024;74(3):229–63.
2. Loud JT, Murphy J. Cancer screening and early detection in the 21st century. Seminars Oncol Nurs. 2017;33(2):121–8.
3. Dlamini Z, Francies FZ, Hull R, Marima R. Artificial intelligence (AI) and big data in cancer and precision oncology. Comput Struct Biotechnol J. 2020;1(18):2300–11.
4. Louis DN, Feldman M, Carter AB, Dighe AS, Pfeifer JD, Bry L, Almeida JS, Saltz J, Braun J, Tomaszewski JE, Gilbertson JR. Computational pathology: a path ahead. Arch Pathol Lab Med. 2016;140(1):41–50.
5. Tsai PC, Lee TH, Kuo KC, Su FY, Lee TL, Marostica E, Ugai T, Zhao M, Lau MC, Väyrynen JP, Giannakis M. Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. Nat Commun. 2023;14(1):2102.
6. Soueidan H, Nikolski M. Machine learning for metagenomics: methods and tools. arXiv preprint arXiv:1510.06621. 2015.
7. Al-Sahaf H, Bi Y, Chen Q, Lensen A, Mei Y, Sun Y, Tran B, Xue B, Zhang M. A survey on evolutionary machine learning. J R Soc N Z. 2019;49(2):205–28.
8. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25(6):954–61.
9. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Med Phys. 2019;29(2):102–27.
10. Luo H, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. Lancet Oncol. 2019;20(12):1645–54.
11. Yap J, Yolland W, Tschandl P. Multimodal skin lesion classification using deep learning. Exp Dermatol. 2018;27(11):1261–7.
12. Ciregan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification. In 2012 IEEE conference on computer vision and pattern recognition. 3642:3649. 2012
13. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. Cham: Springer, Berlin Heidelberg; 2013.
14. Cruz-Roa A, Basavanhally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. Med Imag Dig Pathol. 2014;9041:904103.
15. Ertosun MG, Rubin DL. Automated grading of gliomas using deep learning in digital pathology images a modular approach with ensemble of convolutional neural networks. Am Med Inform Assoc. 2015;2015:1899.
16. Wang H, Cruz-Roa A, Basavanhally A, Gilmore H, Shih N, Feldman M, Tomaszewski J, Gonzalez F, Madabhushi A. Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection. Med Imag Dig Pathol. 2014;9041(66):75.
17. Sirinukunwattana K, Raza SE, Tsang YW, Snead DR, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging. 2016;35(5):1196–206.
18. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans Biomed Eng. 2014;61(6):1729–38.
19. Mallat S. Group invariant scattering. Commun Pure Appl Math. 2012;65(10):1331–98.
20. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, Madabhushi A. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans Med Imaging. 2015;35(1):119–30.
21. Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, Pekmezci M, Schwartzbaum JA, Turner MC, Walsh KM, Wrensch MR. The epidemiology of glioma in adults: a "state of the science" review. Neuro Oncol. 2014;16(7):896–913.

22. Kong J, Cooper LA, Wang F, Gao J, Teodoro G, Scarpace L, Mikkelsen T, Schniederjan MJ, Moreno CS, Saltz JH, Brat DJ. Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates. PLoS ONE. 2013;8(11): e81049.
23. Kong J, Cooper LA, Wang F, Gutman DA, Gao J, Chisolm C, Sharma A, Pan T, Van Meir EG, Kurc TM, Moreno CS. Integrative, multimodal analysis of glioblastoma using TCGA molecular data, pathology images, and clinical outcomes. IEEE Trans Biomed Eng. 2011;58(12):3469–74.
24. Coons SW, Johnson PC, Scheithauer BW, Yates AJ, Pearl DK. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas. Cancer Interdisc Int J Am Cancer Soc. 1997;79(7):1381.
25. Zhang Z, Chen P, McGough M, Xing F, Wang C, Bui M, Xie Y, Sapkota M, Cui L, Dhillon J, Ahmad N. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. Nat Machine Intell. 2019;1(5):236–45.
26. Graham S, Vu QD, Raza SE, Azam A, Tsang YW, Kwak JT, Rajpoot N. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med Image Anal. 2019;1(58): 101563.
27. Fei-Fei L, Deng J, Li K. ImageNet: constructing a large-scale image database. J Vision. 2009;9(8):1037.
28. Arnab A, Miksik O, Torr PH. On the robustness of semantic segmentation models to adversarial attacks. Proc IEEE Conf Comput Vision Pattern Recogn. 2018;888:897.
29. Xu Y, Li Y, Wang Y, Liu M, Fan Y, Lai M, Eric I, Chang C. Gland instance segmentation using deep multichannel neural networks. IEEE Trans Biomed Eng. 2017;64(12):2901–12.
30. Sirinukunwattana K, Pluim JP, Chen H, Qi X, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, Böhm A. Gland segmentation in colon histology images: the glas challenge contest. Med Image Anal. 2017;1(35):489–502.
31. Ren S, He K, Girshick R, Zhang X, Sun J. Object detection networks on convolutional feature maps. IEEE Trans Pattern Anal Mach Intell. 2016;39(7):1476–81.
32. Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: a survey. Med Image Anal. 2021;1(67): 101813.
33. Gong R, Wang L, Wang J, Ge B, Yu H, Shi J. Self-distilled supervised contrastive learning for diagnosis of breast cancers with histopathological images. Comput Biol Med. 2022;1(146): 105641.
34. Abousamra S, Gupta R, Hou L, Batiste R, Zhao T, Shankar A, Rao A, Chen C, Samaras D, Kurc T, Saltz J. Deep learning-based mapping of tumor infiltrating lymphocytes in whole slide images of 23 types of cancer. Front Oncol. 2022;16(11): 806603.
35. Sarker MM, Akram F, Alsharid M, Singh VK, Yasrab R, Elyan E. Efficient breast cancer classification network with dual squeeze and excitation in histopathological images. Diagnostics. 2022;13(1):103.
36. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. Proc IEEE Conf Comput Vision Pattern Recogn. 2016;2424:2433.
37. Fisher R, Perkins S, Walker A, Wolfart E. Spatial filters-laplacian of gaussian. 2003
38. Jia Z, Huang X, Eric I, Chang C, Xu Y. Constrained deep weak supervision for histopathology image segmentation. IEEE Trans Med Imaging. 2017;36(11):2376–88.
39. Liang Q, Nan Y, Coppola G, Zou K, Sun W, Zhang D, Wang Y, Yu G. Weakly supervised biomedical image segmentation by reiterative learning. IEEE J Biomed Health Inform. 2019;23(3):1205–14.
40. Wang S, Zhu Y, Yu L, Chen H, Lin H, Wan X, Fan X, Heng PA. RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. Med Image Anal. 2019;58: 101549.
41. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH, 2015. Efficient multiple instance convolutional neural networks for gigapixel resolution image classification. *arXiv preprint* arXiv:1504.07947.
42. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med. 2019;25(8):1301–9.
43. Wang X, Chen H, Gan C, Lin H, Dou Q, Tsougenis E, Huang Q, Cai M, Heng PA. Weakly supervised deep learning for whole slide lung cancer image analysis. IEEE Trans Cybern. 2019;50(9):3950–62.
44. Qu H, Wu P, Huang Q, Yi J, Riedlinger GM, De S, Metaxas DN, 2019. Weakly supervised deep nuclei segmentation using points annotation in histopathology images. In International Conference on Medical Imaging with Deep Learning (pp. 390–400). PMLR.
45. Li C, Wang X, Liu W, Latecki LJ, Wang B, Huang J. Weakly supervised mitosis detection in breast histopathology images using concentric loss. Med Image Anal. 2019;53:165–78.
46. Yang L, Zhang Y, Zhao Z, Zheng H, Liang P, Ying MT, Ahuja AT, Chen DZ, 2018. Boxnet: Deep learning based biomedical image segmentation using boxes only annotation. *arXiv preprint* arXiv:1806.00593.
47. Chen RJ, Lu MY, Williamson DF, Chen TY, Lipkova J, Noor Z, Shaban M, Shady M, Williams M, Joo B, Mahmood F. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. Cancer Cell. 2022;40(8):865–78.
48. Guo R, Xie K, Pagnucco M, Song Y. SAC-Net: learning with weak and noisy labels in histopathology image segmentation. Med Image Anal. 2023;86: 102790.
49. Zheng Q, Yang R, Xu H, Fan J, Jiao P, Ni X, Yuan J, Wang L, Chen Z, Liu X. A weakly supervised deep learning model and human-machine fusion for accurate grading of renal cell carcinoma from histopathology slides. Cancers. 2023;15(12):3198.
50. Oberlaender M, Dercksen VJ, Egger R, Gensel M, Sakmann B, Hege HC. Automated three-dimensional detection and counting of neuron somata. J Neurosci Methods. 2009;180(1):147–60.
51. Fatakdawala H, Xu J, Basavanhally A, Bhanot G, Ganesan S, Feldman M, Tomaszewski JE, Madabhushi A. Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): application to lymphocyte segmentation on breast cancer histopathology. IEEE Trans Biomed Eng. 2010;57(7):1676–89.
52. Basavanhally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, Bhanot G, Madabhushi A. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. IEEE Trans Biomed Eng. 2009;57(3):642–53.
53. Hu B, Tang Y, Eric I, Chang C, Fan Y, Lai M, Xu Y. Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks. IEEE J Biomed Health Inform. 2018;23(3):1316–28.
54. Bulten, W. and Litjens, G., 2018. Unsupervised prostate cancer detection on H&E using convolutional adversarial autoencoders. *arXiv preprint* arXiv:1804.07098.

55. Sari CT, Gunduz-Demir C. Unsupervised feature extraction via deep learning for histopathological classification of colon tissue images. IEEE Trans Med Imaging. 2018;38(5):1139–49.
56. Sheikh TS, Kim JY, Shim J, Cho M. Unsupervised learning based on multiple descriptors for WSIs diagnosis. Diagnostics. 2022;12(6):1480.
57. Nam JG, Park S, Park CM, Jeon YK, Chung DH, Goo JM, Kim YT, Kim H. Histopathologic basis for a chest CT deep learning survival prediction model in patients with lung adenocarcinoma. Radiology. 2022;305(2):441–51.
58. Niehues JM, Quirke P, West NP, Grabsch HI, van Treeck M, Schirris Y, Veldhuizen GP, Hutchins GG, Richman SD, Foersch S, Brinker TJ. Generalizable biomarker prediction from cancer pathology slides with self-supervised deep learning: a retrospective multi-centric study. Cell Rep Med. 2023. https://doi.org/10.1016/j.xcrm.2023.100980.
59. Hosseini SM. *et al.* (2023) 'Class-imbalanced unsupervised and semi-supervised domain adaptation for Histopathology Images', 2023 45th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC). https://doi.org/10.1109/embc40787.2023.10340049.
60. Fetisov N, Hall L, Goldgof D. and Schabath M, 2023 Unsupervised prostate cancer histopathology image segmentation via meta-learning. In 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS). IEEE.
61. Wang X, Yang S, Zhang J, Wang M, Zhang J, Yang W, Huang J, Han X. Transformer-based unsupervised contrastive learning for histopathological image classification. Med Image Anal. 2022;1(81): 102559.
62. Dippel J, Feulner B, Winterhoff T, Milbich T, Tietz S, Schallenberg S, Dernbach G, Kunft A, Heinke S, Eich ML, Ribbat-Idel J. RudolfV: a foundation model by pathologists for pathologists. arXiv preprint arXiv:2401.04079. 2024 Jan 8.
63. Weiss K, Khoshgoftaar T, Wang D. A survey of transfer learning. J Big Data. 2016. https://doi.org/10.1186/s40537-016-0043-6.
64. Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint* arXiv:1409.1556..
65. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V. and Rabinovich A, 2015. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1–9).
66. He K, Zhang X, Ren S. and Sun J, 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770–778).
67. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M and Adam H, 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint* arXiv:1704.04861.
68. Huang G, Liu Z, Pleiss G, Van Der Maaten L, Weinberger K. Convolutional networks with dense connectivity. IEEE Trans Pattern Anal Machine Intell. 2019;44(12):8704.
69. Tsuneki M, Abe M, Kanavati F. A deep learning model for prostate adenocarcinoma classification in needle biopsy whole-slide images using transfer learning. Diagnostics. 2022;12(3):768.
70. Noorbakhsh J, Farahmand S, Soltanieh-ha M, Namburi S, Zarringhalam K. and Chuang J, 2019. Pan-cancer classifications of tumor histological images using deep learning. *Preprint at* https://doi.org/10.1101/715656v1.
71. Mahmood T, Arsalan M, Owais M, Lee MB, Park KR. Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and deep CNNs. J Clin Med. 2020;9(3):749.
72. Lester SC, Bose S, Chen YY, Connolly JL, de Baca ME, Fitzgibbons PL, Hayes DF, Kleer C, O'Malley FP, Page DL, Smith BL. Protocol for the examination of specimens from patients with invasive carcinoma of the breast. Arch Pathol Lab Med. 2009;133(10):1515–38.
73. Korde LA, Somerfield MR, Carey LA, Crews JR, Denduluri N, Hwang ES, Khan SA, Loibl S, Morris EA, Perez A, Regan MM. Neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast cancer: ASCO guideline. J Clin Oncol. 2021;39(13):1485–505.
74. Balkenhol MC, Tellez D, Vreuls W, Clahsen PC, Pinckaers H, Ciompi F, Bult P, van der Laak JA. Deep learning assisted mitotic counting for breast cancer. Lab Invest. 2019;99(11):1596–606.
75. Pantanowitz L, Hartman D, Qi Y, Cho EY, Suh B, Paeng K, Dhir R, Michelow P, Hazelhurst S, Song SY, Cho SY. Accuracy and efficiency of an artificial intelligence tool when counting breast mitoses. Diagn Pathol. 2020;15(1):1–10.
76. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, Krause J, Niehues JM, Sommer KA, Bankhead P, Kooreman LF. Pan-cancer image-based detection of clinically actionable genetic alterations. Nature Cancer. 2020;1(8):789–99.
77. Figueira G, Wang Y, Sun L, Zhou H, Zhang Q. Adversarial-based domain adaptation networks for unsupervised tumour detection in histopathology. In2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) 2020. 1284–1288. IEEE.
78. Chen RJ, Lu MY, Wang J, Williamson DF, Rodig SJ, Lindeman NI, Mahmood F. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Trans Med Imag. 2020;41(4):757.
79. Saldanha OL, Quirke P, West NP, James JA, Loughrey MB, Grabsch HI, Salto-Tellez M, Alwers E, Cifci D, Ghaffari Laleh N, Seibel T. Swarm learning for decentralized artificial intelligence in cancer histopathology. Nat Med. 2022;3:118.
80. Saldanha OL, Muti HS, Grabsch HI, Langer R, Dislich B, Kohlruss M, Keller G, van Treeck M, Hewitt KJ, Kolbinger FR, Veldhuizen GP. Direct prediction of genetic aberrations from pathology images in gastric cancer with swarm learning. Gastric Cancer. 2023;26(2):264–74.
81. Bhattacharjee S, Kim CH, Prakash D, Park HG, Cho NH, Choi HK. An efficient lightweight CNN and ensemble machine learning classification of prostate tissue using multilevel feature analysis. Appl Sci. 2020;10(22):8013.
82. Lin J-R, et al. Multiplexed 3D Atlas of state transitions and immune interaction in colorectal cancer. Cell. 2023. https://doi.org/10.1016/j.cell.2022.12.028.
83. Bronsert P, Enderle-Ammour K, Bader M, Timme S, Kuehs M, Csanadi A, Kayser G, Kohler I, Bausch D, Hoeppner J, Hopt UT. Cancer cell invasion and EMT marker expression: a three-dimensional study of the human cancer–host interface. J Pathol. 2014;234(3):410–22.
84. Richardson DS, Guan W, Matsumoto K, Pan C, Chung K, Ertürk A, Ueda HR, Lichtman JW. Tissue clearing. Nat Rev Methods Primers. 2021;1(1):84.
85. Chen H, Li C, Wang G, Li X, Rahaman MM, Sun H, Hu W, Li Y, Liu W, Sun C, Ai S. GasHis-Transformer: a multi-scale visual transformer approach for gastric histopathological image detection. Pattern Recogn. 2022;130: 108827.
86. Zhou X, Tang C, Huang P, Tian S, Mercaldo F, Santone A. ASI-DBNet: an adaptive sparse interactive resnet-vision transformer dual-branch network for the grading of brain cancer histopathological images. Interdisc Sci: Comput Life Sci. 2023;15(1):15–31.
87. Vorontsov E, Bozkurt A, Casson A, Shaikovski G, Zelechowski M, Liu S, Severson K, Zimmermann E, Hall J, Tenenholtz N, Fusi N. Virchow: A million-slide digital pathology foundation model. arXiv preprint arXiv:2309.07778. 2023.

88. Chen RJ, Ding T, Lu MY, Williamson DF, Jaume G, Song AH, Chen B, Zhang A, Shao D, Shaban M, Williams M. Towards a general-purpose foundation model for computational pathology. Nat Med. 2024;30(3):850–62.
89. Lu MY, Chen B, Williamson DF, Chen RJ, Zhao M, Chow AK, Ikemura K, Kim A, Pouli D, Patel A, Soliman A. A multimodal generative AI copilot for human pathology. Nature. 2024;634(8033):466–73.
90. Vaidya A, Chen RJ, Williamson DF, Song AH, Jaume G, Yang Y, Hartvigsen T, Dyer EC, Lu MY, Lipkova J, Shaban M. Demographic bias in misdiagnosis by computational pathology models. Nat Med. 2024;30(4):1174–90.
91. Kulinski S, Inouye DI. Towards explaining distribution shifts. Int Conf Machine Learn. 2023;17931:17952.