Journal of Pathology Informatics 14 (2023) 100184

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Review Article

Application of digital pathology and machine learning in the liver, kidney and lung diseases

Benjamin Wu [a,*], Gilbert Moeckel [b]

[a] Horace Mann School, Bronx, NY, USA
[b] Department of Pathology, Yale University School of Medicine, New Haven, CT, USA

ARTICLE INFO

Keywords:
Digital pathology
Machine learning (ML)
Whole slide imaging (WSI)
Liver
Kidney
Lung

ABSTRACT

The development of rapid and accurate Whole Slide Imaging (WSI) has paved the way for the application of Artificial Intelligence (AI) to digital pathology. The availability of WSI in the recent years allowed the rapid development of various AI technologies to blossom. WSI-based digital pathology combined with neural networks can automate arduous and time-consuming tasks of slide evaluation. Machine Learning (ML)-based AI has been demonstrated to outperform pathologists by eliminating inter- and intra-observer subjectivity, obtaining quantitative data from slide images, and extracting hidden image patterns that are relevant to disease subtype and progression. In this review, we outline the functionality of different AI technologies such as neural networks and deep learning and discover how aspects of different diseases make them benefit from the implementation of AI. AI has proven to be valuable in many different organs, with this review focusing on the liver, kidney, and lungs. We also discuss how AI and image analysis not only can grade diseases objectively but also discover aspects of diseases that have prognostic value. In the end, we review the current status of the integration of AI in pathology and share our vision on the future of digital pathology.

Introduction

While the gold-standard of pathology remains pathologists manually reading physical slides and extracting morphology data for diagnosis in both the clinical and non-clinical settings, the interobserver variability limits the comparison of results across studies and sites.[1,2] The discipline has been centered around the light microscope which even into the modern era has remained relatively unchanged.[3] With the acceleration in developments of different fields, especially oncology, pathology modernization is urgently needed to be able to keep up with the increasing demands to create more robust ways of reading, diagnosing, and stratifying patients in a more objective and streamlined manner.[12,13] To this effect, new technologies developed over the last 2 decades have caused the emergence of digital pathology, paving the way for artificial intelligence (AI)-based slide reading and analysis.[5]

One of these technologies is Whole Slide Imaging (WSI), a novel technology that allows for the scanning of microscope slides using a specialized whole slide scanner to create high-resolution images and to digitize histology slides.[1,5] Paired with machine learning (ML) and various types of neural networks, once trained with data, AI can examine digitized microscope slide images and make diagnoses.[9,11] This process is known as image analysis or morphometric analysis. ML is one area of AI, recognizes and learns from a given dataset by algorithms and extrapolates to new dataset for learning with higher accuracy. Deep learning (DL), a particular type of ML, learns data through different layers of artificial neural networks, resembling the function of the neural connection in the human brain.[14] While conventional ML algorithms require more structured data as an input, DL algorithms are able to take in raw data inputs and then employ neural networks in intermediate layers to structure the data and to make it understandable for the other parts of the algorithm.[31–32] DL is particularly effective in pathology as it requires less preparation of complex imaging data associated with slide reading in order for the algorithms to effectively analyze histopathological images.[31]

Important algorithms to assist in ML include various types of neural networks. Neural networks are algorithms that are meant to simulate and mimic the thought processes of a human brain.[6] Neural networks function by taking in different types of data as an input and running them through many hidden layers specified by an algorithm altering the data and eventually returning an output.[7–8] More specifically, different subtypes of neural networks such as deep neural networks (DNNs), convolutional neural networks (CNNs), and artificial neural networks (ANNs) each have their unique usages.[13] Neural network algorithms can extract patterns and learn how different biomarkers or cell structures correspond to different diagnoses. The knowledge learned from the training set is applied to disease

* Corresponding author at: 950 Post Rd., Scarsdale, NY 10583, USA.
E-mail address: benwu200510@gmail.com (B. Wu).

http://dx.doi.org/10.1016/j.jpi.2022.100184
Received 21 September 2022; Received in revised form 28 November 2022; Accepted 28 December 2022
Available online 03 January 2023

diagnosis, patient stratification or prediction of the prognosis based on the WSI slides. As they function, AI will be able to continuously learn from new patient data and refine their models.

This review will discuss how WSI along with ML can be applied to different fields of medicine (Fig. 1). One that we will examine is the use of AI technologies in precision oncology and personalized medicine. Digital pathology along with deep learning has been applied to different cancers in various organs such as the lung, liver, kidney, breast, skin, and whole-body systems to grade diseases objectively and offer accurate prognoses based on ML data. The scope of this review will focus on 3 major organs, the lungs, the liver, and the kidney, and how AI was applied to oncology and non-oncology diseases. This review will highlight different types of cases that benefit the most from AI technologies. Finally, we will look at the future of digital pathology as well as potential drawbacks or regulatory obstacles that need to be overcome in order for digital pathology to become a common practice.[13] (See Tables 1–3).

Whole slide imaging

An essential part of digital pathology is WSI, a novel technology that allows for the scanning and digitization of microscope slides using a specialized high-resolution whole slide scanner.[16,17,28–30] There are multiple vendors making advances in resolution, scanning speed/throughput and robotics to minimize human interaction with the scanner. This generates much higher image resolution and drastically smaller pixel size. Additionally, some manufactures are developing internal QC algorithms to check WSI quality. WSI allows for the automation of previously time-consuming tasks in a high throughput fashion. At the same time, pathologists will have a role in training the AI, compiling datasets, and giving initial diagnoses to train the AI. WSI enables the better organization of microscope slides and a way to both view them digitally and combine them together to form a complete histological image that can be viewed with ease without sacrificing image resolution or being inferior to traditional methods.[18] With the advent of this new technology, slides can now be easily shared between colleagues working at different institutions, something that in the past has been difficult for the community. This also opens up the possibility for large databases or repositories with digitized histological slides to be built that would act as a reference for the pathology community. Furthermore, this technology allows the user to adjust their field of vision to different magnifications, allowing the pathologist to analyze the tissue sample as a whole or magnified on specific cells of interest, similar to that of a traditional light microscope.[3] Finally, while samples on traditional microscope slides degrade over time or slides could be damaged, these digital slides will always show the same structures, and digitizing them allows for less physical storage space necessary.[4] WSI has been proven numerous times to be just as accurate if not more than the traditional light microscopy with regard to diagnostic performance. A study by Mukhopadhyay et al. took pathological specimens from 1992 patients and had 16 surgical pathologists analyze them both under a traditional light microscope and digitally after being scanned by the WSI technology and found that the performance of pathologists when analyzing the slides that had been scanned with the WSI technology was not inferior to that of the traditional light microscopy approach.[3] Therefore, WSI's by themselves are a very powerful tool, allowing pathologists to much easier study and share histological images and examine a large patch of tissue or individual cells with the same scan.

When paired with various ML and DL algorithms, AI will be able to analyze the WSI images, extract disease-relevant information and make diagnoses objectively. For example, Lam et al. employed WSI in the study of esophageal adenocarcinoma where they assessed tissue microarray slides and analyzed the level of staining of tissue samples.[18] Meanwhile, Yuan developed AI methods to analyze the spatial distributions of lymphocytes on triple negative breast cancer WSI scans to help determine the prognosis of patients as well as reactions to chemotherapy and other treatments.[19] Finally, Steele et al. used WSI images from 9 cancer types and AI to enumerate and locate CD8 + tumor-infiltrating lymphocytes and to characterize and organize immunohistochemistry-derived CD8 data in both human and animal tumor samples.[20]

The pandemic helped speed up the adoption of WSI by many pathology labs. However, the regulatory framework for approving the WSI scanners

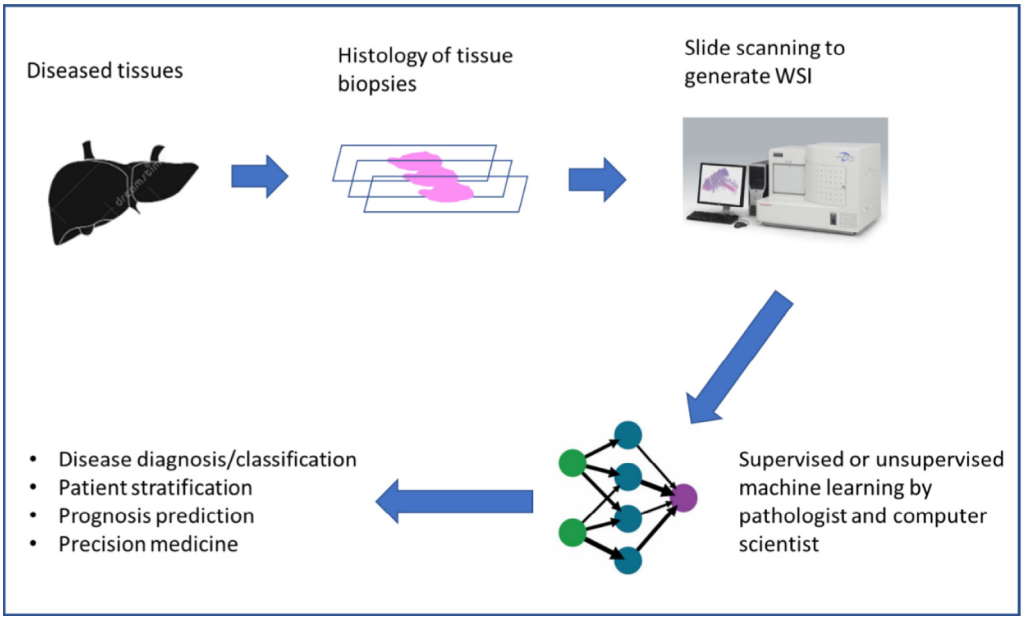

Fig. 1. A simplified workflow of applying digital pathology and machine learning in the clinical setting. Histology slides are made from tissue biopsies of diseased organs and scanned to generate whole slide images (WSIs). WSIs will be segmented for lesions by automated algorithms and go through supervised or unsupervised machine learning based on the quantitative information of key features present in the WSIs by various algorithms. In a supervised learning, algorithms were previously generated using a training WSI set in order to classify the test WSI into certain disease category or grade level. In an unsupervised learning, discriminative features are identified by algorithms without prior labeling. The outcome of the machine learning of features in the WSIs can provide more accurate information to the doctors and scientists for disease diagnosis and classification, which can be used to stratify patients for prognosis and treatment options, achieving precision medicine.

Table 1
A summary of representative studies of the applications of digital pathology in the liver.

Reference	Area of focus	Sample size	Analysis method	Task	Outcome
Jedrzkiewiczl et al.[44]	Liver fibrosis	49	WSI, Photon microscopy	Using AI to standardize fibrosis evaluation methods, making them more accurate and quantitative.	Able to detect the total area of fibrosis and distinguish between collagen fibers and normal tissue.
Masugi et al.[46]	NAFLD	289	WSI, Multivariable logistic regression analysis	Correlate the amount of fibers with fibrosis stage to understand the progression of NAFLD.	Nonlinear relationship between different levels on the Brunt's Fibrosis scale regarding combined collagen and elastin area ratios.
Munsterman et al.[10]	NAFLD	79	WSI, Java plug-in, FIJI	Aimed to quantify steatosis and correlate with NAFLD by calculating the Steatosis Proportionate Area.	Use algorithms to detect steatotic hepatocytes and calculated the Steatosis Proportionate Area. Then correlated this to NAFLD grades.
Native et al.[35]	Macrovesicular steatosis	54	ACM, K-means clustering, Decision tree	Quantify macrovesicular steatosis using segmented liver tissue in histological images.	ACM to amplify important information and extract patterns from the samples. Using important quantitative data, an effective decision tree was created to quantify steatosis.
Schwen et al.	Steatosis in mice	N/A	WSI, 3D analysis	Aimed to quantify zonation of steatosis in mice.	Registered serial sections of mice livers and globule patterns were discerned to determine the different types of lipid vacuoles that were present.
Homeyer et al.[43]	Steatosis	N/A	Image analysis	Fix the unreliability of the current grading system and methods of steatosis.	Used image analysis to separate background and foreground tissue and scored steatosis areas by analyzing lipid blobs.
Atupelage et al.[39]	HCC	109	Five module system	Grade HCC.	Used a five module system that was able to compute textural characteristics, segment nuclei, exclude nuclei that did not belong to hepatocytes and finally grade the HCC.
Kiana et al.[40]	HCC, cholangiocarcinoma	70	WSI, DL, CNN	Distinguish between HCC and cholangiocarcinoma.	Trained a DL model with WSI's and used a CNN in a diagnostic tool to distinguish between the 2 diseases.
Liao et al.[41]	HCC	491	CNN, ML	Distinguish HCC from normal hepatic tissue.	They used a CNN to distinguish between HCC and normal tissue and reused and adapted the CNN to efficiently predict future disease mutations.

remains rigorous and complex. In the United States, the FDA classifies WSI scanners as Class III, indicative of the highest risk.[13]

Machine learning

ML itself is a broad term that includes many different types of algorithms and methods. However, the general principle of ML is that computers are able to learn and improve algorithms with training data and minimal human intervention. The methodology for training the AI can be split into 2 categories, supervised and unsupervised learning.[22] Supervised learning is when the machine learns from labeled data from a training dataset that specifically has a feature that is the target for the AI. For example, histological images may be given to the AI with some slides having specific slide features annotated, including cancer cells, stroma, blood vessels etc. while others being normal matching tissue slides.[23] The AI then will attempt to identify which slides have the cancer and check itself against an answer key. Meanwhile, unsupervised learning is when the AI is given an unbiased dataset and is meant to extract features and patterns by itself. Traditional methods of ML such as hierarchical clustering and K-means clustering are classified as unsupervised learning while algorithms such as neural networks and decision trees are examples of supervised learning. Regardless, ML is effective at image analysis and automating slow processes. An initial investment of training the AI and providing data is required for an efficient method to make slide reading more objective and faster.

Deep learning and neural networks

DL neural networks have been utilized to assist in image analysis of digitized slides. They were first proposed in 1943 by Warren McCulloch and Walter Pitts as a method for an AI to replicate the functions and capability of the human brain.[14] In 1951, the first official neural network was created by Dean Edmunds and Minsky Edmunds.[15] Neural networks have been used in various clinical and non-clinical settings to perform tasks that require some level of cognitive ability.[15] Neural networks employ "neurons", different mathematical functions, that take inputs, passed to them from previous neurons, and put them through specific functions, and then send the

Table 2
A summary of representative studies of the applications of digital pathology in the kidneys.

Reference	Area of focus	Sample size	Analysis method	Explanation	Outcome
Simon et al.[48]	Glomerular detection	N/A	WSI, SVM, LBP, Deep CNN	LBP was adapted for the use of glomerular detection and trained a SVM and Deep CNN to recognize glomeruli.	The algorithms and ML methods were able to detect glomeruli in renal tissue with high accuracy (3% false positive) and can be applied to numerous species such as mice, rats, and humans.
Rosenberg et al.[50]	Enumerating glomeruli	277	WSI, Annotation algorithm	Used WSI's to enumerate glomeruli and compared them to the performance of pathologists using traditional light microscopy.	It was discovered in some instances that the ratio of the amount of glomeruli counted by the algorithm to the amount by the pathologists was 2:1, showing the large inaccuracy of manual enumeration.
Furness et al.[55]	Acute renal transplant rejection	100	Neural networks	100 biopsies showing different signs of renal rejection, some more difficult to categorize than others, were used to train the neural networks to recognize transplant rejection.	The neural networks were able to outperform pathologists. In a set of 21 more difficult to diagnose slides, the neural network recognized 19/21, 1 more than the best performing pathologist.
Yeh et al.[54]	ccRCC	39	WSI, SVM	An SVM was utilized to automatically detect nuclei in the issue sample and to count the amount of nuclei within the sample. Kernel regression was also used to estimate the spatial distribution of nuclei.	They found that the spatial distribution of tumors calculated by the SVM was able to locate areas of necrosis or large nuclei which in turn gave prognostic value.
Tian et al.[52]	ccRCC	395	WSI, ML, Lasso regression	From each WSI, 5 ROI's were graded by both the pathologists and the lasso regression model to separate the ccRCC cases into high and low on an altered Fuhrman grading scale.	The grades assigned by the ML algorithm had prognostic power while the manual ones did not. The high low grading system was much simpler and had an accuracy of (83%).

Table 3
A summary of representative studies of the applications of digital pathology in the lungs.

Reference	Area of focus	Sample size	Analysis method	Explanation	Outcome
Teramoto et al.[62]	LUAD, LUSC, Small cell carcinoma	76	Deep CNN	Use deep CNN's to analyze cytological images and classify between different types of lung cancers	The deep CNN, when analyzing the microscopic images had an accuracy of 70%, which proves that deep CNN's are indeed useful in this type of image classification
Yu et al.[65]	LUAD, LUSC	884	WSI, CNN	Use a CNN to classify different types of NSCLC's and evaluated the performances by measuring AUC.	The CNN was able to outperform the pathologists, with the CNN having an AUC > 0.935 and the pathologists having an AUC > 0.877.
Sha et al.[66]	NSCLC	130	WSI, DL	Predict PD-L1 status of NSCLC samples and find the correlation between PD-L1 status and tumor patterns.	The model was able to accurately predict PD-L1 status when given the test cohort (n = 82).
Hou et al.[61]	LUAD, LUSC	539	WSI, CNN	Used a patch-based CNN to classify between different subtypes of NSCLC, including more difficult ones. The patch was able to output a cancer type probability which was turned into a final diagnosis.	Overcame difficulties of tumor heterogeneity by using MIL to group together different instances that the machine could then learn from.
Coudray et al.[26]	LUAD, LUSC	1634	WSI, CNN	Developed 2 CNN's, 1 to classify between a certain NSCLC subtype and healthy tissue and another to predict mutating genes related to LUAD.	Their first CNN had an AUC of 0.97 when distinguishing between LUAD and normal tissue and 0.99 between LUSC and normal tissue, higher than previous studies (AUC < 0.9) Their second CNN was able to stratify patients based on different genetic.
Yu et al.[58]	NSCLC	2480	Image segmentation pipeline, net-Cox proportional hazards model	Use ML techniques to analyze samples of NSCLC and extract features that could in turn predict survival outcomes.	Found that features such as nuclei decomposition and texture had significant prognostic value.
Wang et al.[56]	LUAD	389	Deep CNN	Use a deep CNN model to recognize and diagnose LUAD.	Found that important features such as shape, features, and structures of tumors had prognostic value. Their deep CNN was successful in recognizing the correct tissue 89.8% of the time.
Saltz et al.[67]	TME, TIL	5202	WSI, CNN	Map TILs and the tumor microenvironment using a CNN that could perform computational staining and analysis for TILs it the given tissue sample.	They were able to use the gathered information to calculate a spatial fraction of TILs. They also found that different TIL patterns and structures correlated to different survival outcomes.
Yi et al.	Microvessel counting	350	FCN	Use and FCN to predict microvessels in lung images.	Used the information to calculate the area of the microvessels and the amount of cells in proximity to them. They showed that there was a positive relationship between the density of microvessels and positive survival outcomes.
Wang et al.[71]	LUAD	1914	CNN	They used a CNN to identify tumor, stromal, and lymphocytes in images of LUAD and then created a spatial map out of these features.	The spatial maps were used to create a prognostic model. They were able to conclude that a higher abundance of stromal cells correlated to a better prognosis, though this model does not take into account different subtypes of lymphocytes

output to other neurons in different layers of the neural networks. Neural networks are extremely effective as they do not require engineered features to translate the primary data into something meaningful but rather can learn directly from it.[21] In order for neural networks to be effective, they must first be trained with a dataset. The datasets are typically image-based so that algorithms can be created to diagnose patients based on the morphology and structure in the tissue images.

There are various types of neural networks and each has its specific uses. The most popular and widely used neural network is the convolutional neural network (CNN).[13] A CNN is a model that excels at processing data that has a grid pattern, such as digitized slides, and can learn spatial hierarchies of features and independently extract patterns.[24] A CNN usually is split into 3 types of layers: convolution, pooling, and a fully connected layer. Convolution and pooling are able to extract features and patterns from the input. Then the fully connected layer translates the extracted features into the form of an output. CNN's are independent and do not require the separate extraction of hand-crafted features within the image. The layers of a CNN are not fully connected to each other. This means that the neurons in one layer are only connected to other layers at specific points and with limited neurons. The goal of a CNN is to analyze a complex image and break it down into the core components such as different shapes, curves, lines, and planes (if the scan is 3-dimensional) and identify useful features for diagnosis.[13] CNN's have been used in many different areas, especially oncology, to provide diagnoses and gradings that are on par with board-certified pathologists. They also help to automate arduous tasks and unlock new and more efficient ways for diagnosing patients or analyzing tumors. Chang et al. used residual CNN's to predict the isocitrate dehydrogenase status of gliomas in different patients to help in prognosis and to help guide the course of medical treatment.[25] Meanwhile, Bejnordi et al. employed Stacked CNN's (feeding one CNN's output into another as an input) to examine histopathological tissue scans of breast lesions and aimed to use the CNN's to categorize breast WSI scans into benign or malignant, namely whether the scans showed ductal carcinoma in-situ or invasive ductal carcinoma.

Liver

WSI technologies and machine learning have seen effective applications in histological analysis of hepatic tissue to help evaluate and treat various diseases such as Non-Alcoholic Fatty Liver Disease (NAFLD), steatosis, and hepatocellular carcinoma. There has also been an interest in how AI can assist in liver transplants.[29]

Non-oncology

Quantification of fibrosis and steatosis

AI-based imaging analysis of WSI holds the promise to deliver objective and quantitative histopathology evaluation and thereby increase precision. It helps move away from inter-observer variability ensuring that appropriate and uniform diagnoses are given. Fibrosis grading is crucial in understanding the disease stage and creating treatment plans. AI has helped standardize evaluation methods for fibrosis, making it more quantitative and accurate.[44–47,81–83] Jedrzkiewiczl et al. quantified liver fibrosis from liver biopsies of various liver diseases using both WSI's and photon microscopy.[44] They used photon microscopy combined with second harmonic generation analysis (the method by which 2 photons can combine to double their frequency and half their wavelength) to distinguish the

collagen fibers from normal tissue. They demonstrated the utility of the WSI and AI in the objective quantification of fibrosis in the clinical liver disease samples and further proposed that these technologies have the resolution to analyze the spatial distribution of the collagen fibers. In another study aiming to determine the relationship of the quantity of fiber and the fibrosis stage, an automated computational method was applied to quantify collagen and elastin fibers in WSIs of Elastica van Gieson-stained liver biopsy specimens from 289 NAFLD patients across multiple hospitals.[46] This study revealed a nonlinear relationship between the combined collagen and elastin area ratios, with stage 4 fibrosis graded on Brunt's fibrosis scale containing a far greater number of fibers compared to stages 0–3.

Accurate evaluation of steatosis is essential to the grading of liver diseases and WSI-based digital analysis has been applied for quantitative assessment of steatosis in H&E slides. Munsterman et al. aimed to quantify steatosis using WSI and digital image analysis to grade disease progression of NAFLD.[10] They used an automated algorithm implemented in Fiji, an open-source biological image analysis platform[33] to grade individual patches of the slide.[10] The algorithm detected the areas of steatosis on the WSI of the H&E slide by a size and roundness-based classifier and then calculated the steatosis proportionate area (SPA) per WSI. There was a strong correlation between SPA and the traditional steatosis grade, demonstrating that AI can be used to objectively quantify steatosis in an automated fashion, a clear unmet need by manual reading. Similarly, Nativ et al. applied automated digital image analysis to segmented liver tissue structures to quantify large droplet macrovesicular steatosis in order to determine liver transplantability.[35] An Active Contour Model (ACM) was used to improve the differentiation of large droplets from small droplets by examining cell nuclei displacement and droplet size information. The results showed an accurate separation of the large lipid droplets and small lipid droplets with a 93.7% specificity and 99.3% sensitivity and a good correlation ($R^2 = 0.97$) with the manual assessment by pathologist.

In a non-clinical setting, novel methods that incorporate image analysis into new and objective ways of scoring steatosis in mice livers have proven useful.[42,43] The liver is a highly structured organ and hepatic physiological and metabolic processes, as well as pathological conditions, display spatial heterogeneity. Liver biopsies represent only a small part of the liver rather than the whole organ. In order to understand the interaction of pathological conditions with metabolic processes in different zones of the liver and detect steatosis in the whole organ, Schwen et al. registered serial sections of WSI of the entire mouse liver.[42] The registered images were combined by AI to obtain 3D analyses for the purpose of quantifying zonation of steatosis. The globule patterns were discerned using different stainings and the types of lipid vacuoles that were present were determined. The results reiterated that quantitative zonated assessment by AI complements the visual description by pathologists. Meanwhile, Homeyer et al. attempted to overcome the spatial heterogeneity of steatosis and reliably detect the steatosis in the liver histological images.[43] Focused score was developed using tile-based hotspot analysis to objectively calculate statistics. Using focused score analysis, they were able to discriminate steatosis of different sizes with good accuracy.

Prior and post liver transplant assessment

Liver transplant has also seen the effective use of WSI and AI.[36,84,85] Frequent biopsies and examinations of histological slides are required for transplants, including confirmation of the diagnosis and the health of the donor liver.[36,84] Post-transplant, the allograft must be closely monitored to ensure the success and acceptance of the donor liver. These tasks would benefit greatly from digital image analysis as it would speed up slow processes and create more inter-observer concordance.[36,85] Furthermore, digital pathology promises to offer the advantages of improved diagnostic reproducibility, identification, and quantification in this highly specific field.[36,85]

Oncology

HCC diagnosis and grading

Hepatocellular carcinoma (HCC) is generally graded on the Edmonson-Steiner system, which has 4 different grades, G1 being the most differentiated and G4 the least. The grades are important as they are the basis for treatment steps and prognosis.[38] This has motivated many researchers to experiment with using digital analysis to grade liver biopsies.[39] Atupelage et al. used a 5-module method to identify liver cells and grade digitized slides of HCC.[39] The module steps compute pixel-wise textural characteristics, segment the nuclei of all of the cells in the image, exclude the non-hepatocyte nuclei belonging to the lymphocytes, histiocytes, and fibroblasts, and finally grade HCC according to a defined HCC grading framework.[39] This algorithm was able to classify HCC into 5 grades (normal liver, G1–G4 HCC) with a high correct classification rate (95.97%). Hepatocyte nuclear texture was identified as the major feature for grading HCC and the accuracy increased with additional features.

Liao et al. aimed to use AI to diagnose HCC and predict the underlying mutations based on exhibited pathology phenotypes.[41] First, a classification CNN was trained using WSIs of HCC and matching normal adjacent tissue samples from The Cancer Genome Atlas (TCGA) and HCC tissue microarrays from The Biobank of West China Hospital (WCH). The algorithm was able to distinguish HCC from normal tissue in an automated fashion with high accuracy. Next, a layer of the classification CNN was modified to extract predictive pathology features from a training set of WSIs from the TCGA to predict single gene mutation. The model was able to predict mutation status for CTNNB1, p53, CSMD3, and ALB with high predictive values. However, the predictive values changed drastically for different genes when using the test set from the TCGA or the external validation set from the WCH, suggesting differences or bias exist between samples in the 2 databases. This example highlights the importance of building robust databases using a standardized sample collection, selection, and digitization procedure in order to maximize and generalize the power of AI models. Nevertheless, this study demonstrates the potential utility of AI-based digital pathology, combined with the genomic data in the diagnosis and subtyping of HCC for precision medicine.

Kiana et al. focused on creating a DL method to help pathologists distinguish between types of similar hepatic cancers, specifically between HCC and cholangiocarcinoma.[40] They employed a cloud-deployed DL model and trained it using a dataset of WSIs of HCC and cholangiocarcinoma. They also used a CNN in a diagnostic tool where WSI scans could be uploaded and the AI would then give its diagnosis and recommendations.

In another study, Feature Aligned Multi-Scale Convolutional Network was employed to improve the detection performance of HCC based on WSI.[86] The algorithm references neighboring information in the whole slide and outperforms the performance of the Single-Scale Convolutional Network.[86]

Kidney

Non-oncology

To streamline the pathology review for disease classification in clinical trials and minimize inter- and intra-observer subjectivity, the Nephrotic Syndrome Study Network (NEPTUNE),[50,51] the first multicenter consortium was created to provide a central digital pathology repository of WSI digitized kidney biopsies. Glass slides were scanned and uploaded at local or central scanning sites to a central database and annotated and graded. NEPTUNE has developed its own methods for scoring and grading to promote inter-observer reproducibility. The creation of NEPTUNE has standardized methods of grading slides and promotes the cooperation of pathologists. NEPTUNE has also served as a comprehensive database for training AI as the pre-scored and annotated slides allow for its application in supervised learning.

Segmentation and enumeration of glomeruli for diagnosis and prognosis

In nephropathology, glomeruli localization and quantitative enumeration are fundamental to the accurate diagnosis of renal diseases and treatment.[50] To this aspect, AI-based automated high throughput methods using WSIs have demonstrated prominent advantages over traditional manual reads.[48,50,62] Rosenberg et al. used the biopsy WSIs from the NEPTUNE

repository to count the glomeruli and determine the percent of globally sclerotic glomeruli (GS) by manual enumerating annotated serial digitized tissue sections.[50] They found that the ratio of the number of glomeruli counted by pathologists annotated WSI's to the number counted when using the traditional microscope could be as much as 2:1. There were limits to conventional light microscopy methods when pathologists had to enumerate glomeruli across multiple slides and tissue sections. The percent of GS was also significantly underestimated by conventional enumeration using light microscopy when the percent GS was over 40%. These results suggest that the glomeruli assessment using a computer-aided combination of multiple WSI sections will generate more accurate data for disease classification and subsequent effective treatment and the identification of prognostic markers. This study also demonstrated that even when simply changing observation methods from a light microscope to WSI digitized slides (without image analysis algorithms), accuracy was greatly improved.

In a study conducted by Simon et al.,[48] Local Binary Pattern (LBP), a textural and image analysis method, was adapted for glomeruli detection in WSI digitized slides. The LBP was used to train a Support Vector Machine (SVM) Model and was paired with a Deep CNN trained to recognize glomeruli. The algorithms were able to detect glomeruli in surrounding renal tissue of the biopsy with high precision (false-positive percentages below 3%), offering robust adaptability to a variety of staining methods in multiple species including rats, mice, and humans.

Kidney allograft

The application of AI to detect features such as inflammation, fibrosis, and tubular atrophy in renal allograft is reviewed by Farris et al.[78,79] Herman et al. used CNNs to detect and quantitate inflammatory infiltrates in the biopsies of kidney transplant and demonstrated correlation with interstitial fibrosis and tubular atrophy.[80]

It is often difficult for conventional methods to diagnose early stages of allograft rejection, a stage that is critical in preventing further damage. Furness et al. employed neural networks to analyze biopsies and predict acute renal transplant rejection.[55] They trained their neural network using data and graft biopsies that had been graded by numerous pathologists within the United Kingdom to simplify the supervised learning process. The training set included 100 standard slides and a group of 25 slides that were difficult to classify, being borderline between showing signs of rejection and a healthy transplant. Their neutral network demonstrated high performance for both regular and difficult cases, correctly diagnosed all 20 standard cases and 19 out of 21 difficult ones, 1 more than the best performance by any pathologist in this study. By employing the more difficult training set, they were able to achieve their goal of having the neural network outperform the pathologists in areas where there were traditionally struggles.

Oncology

Grading and prognosis of clear cell renal cell carcinoma

Histologic subtype of kidney tumors correlates with prognosis and patient survival.[63] Clear cell renal cell carcinoma (ccRCC) represents the most common type of malignant tumor in the kidney with a 5-year survival of 76%.[63] Although the traditional 4-tiered Fuhrman nuclear grading system is still the gold-standard for ccRCC, there is an innate level of moderate inter- and intra-observer variability due to the complicated nature of the 4 tiers.[53]

WSI's and image analysis methods were first explored by Yeh et al. to automatically grade ccRCC.[54] Thirty-nine digitized H&E stained ccRCC slides of varying grades were analyzed by automatic stain recognition algorithms. An SVM classifier was trained to automatically isolate and detect nuclei in the tissue sample. First, different slides were taken and segmented into smaller sizes. In the next step, SVM counted the number of nuclei in each of the smaller pieces of the grid. Finally, the sizes of the nuclei were estimated and kernel regression was used to estimate the spatial distribution of the nuclei to determine the grade of the ccRCC. They found that on the Furhman nuclear grading scale, grades 1 and 2 were nearly indistinguishable as with 3 and 4. As a result, they decided to simplify the grading into high or low. The spatial distribution of the nuclei can also be used to locate regions of the tumor with larger nuclei or with necrosis which are very important in helping to determine a prognosis for patients.

In a recent study, Tian et al. used a large sample set to demonstrate the power of WSI and ML in improving ccRCC grading systems with prognostic significance. The WSI images of 395 ccRCC cases along with the clinical data were used for automated analysis to develop a 2-tiered Fuhrman's grading system.[52] Five regions of interest (ROIs) from each WSI were picked and graded manually by pathologists according to the Fuhrman's grading system as well as by an automated computer-based imaging analysis and a Lasso regression trained ML model. The Lasso model split the grading into 2 tiers: high and low, which simplified the system while providing high accuracy (83%). There was a significant association between predicted grade and overall survival thereby providing prognostic value. High-grade cases predicted by the Lasso model were associated with poorer prognosis and lower overall survival rates compared to the low-grade cases. In contrast, the grades assigned by manual reading do not have prognostic power.

Lung

The applications of digital pathology and AI in the lung have been almost exclusively limited to the clinical oncology setting. Despite the success in oncology, there has been limited application of AI towards the non-oncology lung diseases. One non-oncology paper published recently described combining human expertise and deep learning algorithm to developed models to extract features to diagnose interstitial pneumonia with high accuracy.[87]

Diagnosis and classifications of lung cancers

The diagnosis of lung cancer subtypes and stages relies on the examination of the morphological phenotype of lung biopsies, a challenging and labor-intensive task for even highly experienced pulmonary pathologists. The accurate classification of lung cancer subtypes and stages is crucial in determining the appropriate therapy out of the many treatment options (chemotherapy, targeted therapy, or immunotherapy) as well as predicting a patient's prognostic outcome. However, the current gold-standard for diagnosing and grading lung cancers is highly subjective and qualitative and not able to capture the tumor heterogeneity and existence of multiple different subtypes in the same patient. Conceptually, WSI of lung biopsies coupled with AI offers potential advantages in the diagnosis by leveraging the rich collection of tumor samples of diverse subtypes and stages and matched normal samples in a centralized lung cancer repository. The power of DL can also be harnessed to identify key features and train the model to recognize these features in a fully automated fashion. To this aspect, AI and specifically CNN's have been used extensively to help diagnose and classify different subtypes of lung cancer, especially for the 2 most prevalent subtypes, adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC).[26,61,62,65,66] Hou et al. used a patch-based CNN to classify subtypes of NSCLCs including mixed-subtypes which are particularly difficult to handle.[61] They extracted pathological features and trained classifiers to output a cancer type probability at a patch level and then aggregate the patch-level classification into a final decision of the entire image. They also utilized an expectation maximization model to identify distinct and important patches within the WSI and feed them to the CNN as training data so that the cancer types would be recognized solely by important and distinct areas. To overcome traditional CNN's difficulty in dealing with tumor heterogeneity, they used Multiple Instance Learning (MIL) to allow a more efficient supervised learning that is able to group together different instances that the machine can then learn from.

In a comprehensive study, Coudray et al. similarly developed deep CNN algorithms using 1634 WSIs from TCGA to classify LUAD, LUSC, and normal lung tissues independently collected at their own institution.[26] They also trained and tested their models using pixel tiles and aggregated the tile-level classification by averaging the prediction probability into a

slide-level classification. The accuracy (AUCs of 0.97 and 0.99 for LUAD and LUSC classification and tumor versus normal tissues, respectively) was significantly higher than previous studies (AUC < 0.9) and on par with the accuracy of pathologists. Furthermore, using WSIs taken from the Genomic Data Commons Database, they trained another CNN model to predict frequently mutated genes pertaining to LUAD and demonstrated that mutations of 6 genes (STK11, EGFR, FAT1, SETBP1, KRAS, and TP53) can be predicted by their model using pathology images alone with reasonable accuracy. These promising results suggest that the deep CNN can not only assist pathologists to diagnose and classify LUAD, but also can stratify patients based on their genetic mutation. As the treatment therapies for LUAD and LUSC differ significantly, gene mutation can further affect the choice of therapies. This information is crucial in determining the most appropriate treatment and subsequently achieving precision medicine.

Cancer prognosis

The current gold-standard for diagnosing and grading is not able to determine disease progression and predict survival outcomes.[59] Another promising application of AI in lung oncology is to explore the relationship between pathological features, patient prognosis, and survival outcomes. Both Yu et al. and Wang et al. aimed to use ML to analyze NSCLC tumor samples to extract key features and link them to the survival outcomes.[56,58] Yu et al. built fully automated ML models based on 2186 WSIs of LUAD and LUSC from TCGA and successfully distinguished shorter- and longer-term survivors of Stage I LUAD and LUSC.[58] First, they used an image segmentation pipeline, where a series of connected algorithms are used to segment images and extract quantitative tumor features. The top classifier features were selected to train the models to distinguish the survival outcomes. Models were then validated with 294 images from the Stanford Tissue Microarray Database. After testing the effectiveness through a net-Cox proportional hazards model, they found that features such as nuclei decomposition and texture had significant prognostic value and helped predict survival outcomes.[58,60] Meanwhile, Wang et al. used a deep CNN model to recognize tumors and diagnose LUAD. They found that tumor structure, features, and shapes could have prognostic value.[56] They trained an automated deep CNN model to detect tumor regions, recognizing correct types of tissue 89.8% of the time. They then developed the prognostic model that could determine survival rate using tumor regional shape. A model was built using the National Lung Screening Trial (NLST) cohort and validated independently by a TCGA LUAD cohort. Their model, based on 15 well-defined tumor features such as area, perimeter, filled area, and convex area, was able to categorize people into low- and high risk. After adjusting for affecting patient demographics such as age and gender, the risk model independently predicted survival outcomes with high statistical significance. These novel modeling approaches by Yu and Wang have demonstrated that quantitative pathological features can be utilized to successfully predict the disease prognosis and survival outcomes of LUAD and LUSC patients and their accuracy is superior to the prognostic prediction based on the tumor grade and stage calls made by the conventional manual pathology reading. One can expect that similar approaches can be expanded to prognosis prediction for other tumor types.

In a recent study, Qaiser et al. used a weakly supervised survival convolutional neural network (WSS-CNN) with a visual attention mechanism to predict the overall survival of the lung and bladder urothelial carcinoma.[88] They demonstrated that the features identified is predictive of clinical outcomes of both tumors and could be used to stratify patient for precision medicine.

Tumor microenvironment (TME) characterization

The TME is a complex and heterogeneous system consisting of an extracellular matrix, stromal cells, and immune cells and it plays an important role in supporting tumor growth and tumor invasion from immune surveillance.[73] Characterization of TME can inform treatment decisions and predict patient responses. Tumor-infiltrating lymphocytes (TIL) are generally a positive prognostic factor and their spatial organization in the TME may be a key factor in the potential response to immunotherapies.[74] Characterization of TME, specifically the study of TILs, has also seen the application of neural networks to the analysis for prognostic value.[57,68,69] The approach of these studies is to map the TILs within the lung tissue using neural networks and then correlate to survival. Saltz et al. accomplished this by using a CNN that was able to perform computational staining and analysis for the presence of TILs in patches of WSI images from 13 tumor types available in TCGA.[67] A spatial fraction of TILs was then calculated and local spatial structures of TILs were assessed. Statistics analyses showed that different TIL patch patterns and structures correlated to the survival outcomes.

Other microenvironment aspects also hold prognostic potential. Microvessels and angiogenesis is an important target in the treatment of a malignancy.[70] Traditional microvessel counting is labor-intensive and difficult to perform manually, however this information is crucial and has not been used to its potential due to feasibility issues holding it back. Yi et al. used fully convolutional networks (a type of neural network) to predict microvessels in H&E-stained lung images.[70] They then calculated both the amount of area that the microvessels occupied and the percentage of cells in proximity to the microvessels that were tumor cells. There was a strong correlation between the density of microvessels and a positive survival outcome. Similarly, Rączkowski et al. applied deep learning to segment TME of NSCLC and demonstrated predictive power of tumor mutations and patient survival.[89]

Mapping different types of cells in the TME is impractical by manual reading and would benefit greatly from automation. Wang et al. used a CNN to accurately identify stromal cells, tumor cells, and lymphocytes in the pathological images of LUAD, and essentially create a “spatial map” of 3 cell types of tumor, stromal, and lymphocyte cells out of the pathology image.[71] Image features extracted from these “spatial maps” were used to develop and predict a prognostic model. The statistical analysis showed that high abundance of stromal cells correlated with a better prognosis. This result has to be taken with the consideration that their algorithm only recognizes 3 cell types and is incapable of discerning lymphocyte subtypes without the immunostaining. Despite these limitations, the study serves as a starting point in the application of AI in TME characterization, highlights the challenges in quantitative characterization of the complex TME, and calls for further advances in this important area of research.

Overcoming hurdles

AI has the potential to revolutionize pathology and there is a growing interest in applying WSI and AI to pathology evaluation in both the clinical and non-clinical research. Though it has seen success, it currently has not reached the state of being a standard part of the discipline. The use of AI requires both extensive monetary and time investment. In order for WSIs to be effective, they often must be paired with algorithms that are able to perform image analysis. It takes a considerable amount of expertise and experimentation to develop and validate the AI algorithms before they can be used to analyze datasets and extract meaningful information. Pathologists would also have to be trained to work with the AI. Furthermore, AI technologies are not simple to market. In fact, the United States FDA rates AI-based software as Risk level II or III: the high-risk grades which requires a minimal 510(k) approval pathway to market. In other countries too, such as ones in the European Union, there are strict laws on the use of AI technologies and producers face similar restricting regulations in those countries as in the United States.[76,77] This has caused significant hurdles to market and proliferate WSI scanners and other AI devices.

Despite these challenges, this field of research is rapidly evolving and there is growing interest for doctors, pathologists, engineers, and computer scientists to collaborate.[6] One successful example of such collaboration is the creation of the aforementioned kidney database NEPTUNE.[51] NEPTUNE has developed intricate methods so that digitized slides can be uploaded locally or at centralized locations and has even developed their own scoring systems for various diseases to promote inter- and intra-

observer concordance and uniformity. Meanwhile, the Digital Pathology Association has also compiled WSIs of various tissues into a comprehensive repository that can be accessed through their website. Still other organizations, which do not specifically specialize in WSI, have their own sizable repositories, with their data being frequently used in different studies and clinical trials.[26] The largest of these being TCGA. The Innovative Medicines Initiative (IMI), a consortium between the European Union is building BIGPICTURE, a central repository of digital pathology slides to support the development of AI tools.[75] The Pharmaceutical Industry Consortium plans to contribute 3 million high quality digital pathology slides from non-clinical safety studies and clinical studies. Digital slides from clinical trials will also be contributed. Additionally, the consortium member companies will provide data harmonization expertise and guidance on interactions with health authorities. This initiative is expected to bring tremendous advancement in the integration of digital pathology and AI in non-clinical and clinical settings. Groups such as Pantanowitz et al. are working towards testing the capabilities of AI and WSI and validating their usage as diagnostic tools so that they can be incorporated seamlessly into pathology workflows.[72]

Conclusions

Digital pathology paired with ML has demonstrated substantial advantages over the conventional labor-intensive, low-throughput, qualitative, and descriptive manual slide reading by pathologists in the diagnosis, classification, and prognosis of various diseases. The advent of WSI technologies has brought opportunities to build centralized image repositories and collaborate between institutions around the world. In the meantime, it has also brought challenges of how to process the vast amount of the digital image data and extract meaningful information from it. Deep ML-based AI has the power of processing large datasets, extracting features and patterns that are not visible to the human eye, and making more informed decisions regarding disease treatment and patient management. As the technologies develop and the regulatory pathways clear, an extensive integration of AI in pathology evaluation in clinical and non-clinical settings is expected to occur in the next decade. Collectively, these events will transform the current practice by providing objective and deeper knowledge of digital images, helping pathologists and doctors make informed decisions, thereby achieving precision medicine and ultimately improving patient's health.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Shanes JG, Ghali J, Billingham ME. Ferrans V.J., Fenoglio J.J., Edwards W.D., et al. Interobserver variability in the pathologic interpretation of endomyocardial biopsy results. Circulation 1987;75(2):401–405. https://doi.org/10.1161/01.cir.75.2.401.
2. Potts SJ, Young GD, Voelker FA. The role and impact of quantitative discovery pathology. Drug Discov Today 2010 Nov;15(21–22):943–950. https://doi.org/10.1016/ j.drudis.2010.09.001. Epub 2010 Oct 12.
3. Mukhopadhyay S, Feldman MD, Abels E, Ashfaq R, Beltaifa S. Cacciabeve N.G., et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). Am J Surg Pathol 2018;42(1):39–52. https://doi.org/10.1097/PAS.0000000000000948.
4. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. Med Image Anal 2016;33:170–175. https://doi.org/10.1016/j.media.2016.06.037.
5. Pantanowitz L, Sharma A, Carter AB, Kurc T, Sussman A, Saltz J. Twenty ye ars of digital pathology: an overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. J Pathol Inform 2018;9:40. Published 2018 Nov 21: https://doi.org/10.4103/jpi.jpi_69_18.
6. Cohen S, Furie MB. Artificial intelligence and pathobiology join forces. Am J Pathol 2019;189(1):4–5. https://doi.org/10.1016/j.ajpath.2018.11.002.
7. Araújo T., Aresta G., Castro E., Rouco J., Aguiar P., Eloy C., et al. Classification of breast cancer histology images using convolutional neural networksPLoS One.2017126e0177544 Published 201
8. Kriegeskorte N, Golan T. Neural network models and deep learning. Curr Biol 2019;29(7):R231–R236. https://doi.org/10.1016/j.cub.2019.02.034.
9. Komura D, Ishikawa S. Machine learning approaches for pathologic diagnosis. Virchows Arch 2019;475(2):131–138. https://doi.org/10.1007/s00428-019-02594-w.
10. Munsterman ID, van Erp M, Weijers G, Bronkhorst C, de Korte CL. Drenth J.P.H., et al. A novel automatic digital algorithm that accurately quantifies steatosis in NAFLD on histopathological whole-slide images. Cytometry B Clin Cytom 2019;96(6):521–528. https://doi.org/10.1002/cyto.b.21790.
11. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB. Contis L., et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med 2013;137(12):1710–1722. https://doi.org/10.5858/arpa.2013-0093-CP.
12. Cornish TC, Swapp RE, Kaplan KJ. Whole-slide imaging: routine pathologic diagnosis. Adv Anat Pathol 2012;19(3):152–159. https://doi.org/10.1097/PAP.0b013e318253459e.
13. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16(11):703–715. https://doi.org/10.1038/s41571-019-0252-y.
14. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. 1943. Bull Math Biol 1990;52(1–2):97–99.
15. Park WJ, Park JB. History and application of artificial neural networks in dentistry. Eur J Dent 2018;12(4):594–601. https://doi.org/10.4103/ejd.ejd_325_18.
16. Rojo MG, García GB, Mateos CP, García JG, Vicente MC. Critical comparison of 31 commercially available digital slide systems in pathology. Int J Surg Pathol 2006;14(4):285–305. https://doi.org/10.1177/1066896906292274.
17. Yagi Y, Gilbertson JR. A relationship between slide quality and image quality in whole slide imaging (WSI). Diagn Pathol 2008;3 Suppl 1(suppl 1):S12. Published 2008 Jul 15: https://doi.org/10.1186/1746-1596-3-S1-S12.
18. Lam AK, Leung M. Whole-slide imaging for esophageal adenocarcinoma. Methods Mol Biol 2018;1756:135–142. https://doi.org/10.1007/978-1-4939-7734-5_12.
19. Yuan Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. J R Soc Interface 2015;12(103):20141153. https://doi.org/10.1098/rsif.2014.1153.
20. Steele K.E., Tan T.H., Korn R., Dacosta K., Brown C., Kuziora M., et al. Measuring multiple parameters of CD8+ tumor-infiltrating lymphocytes in human cancers by image analysis. J Immunother Cancer. 2018;6(1):20. Published 2018 Mar 6 https://doi.org/10.1186/s40425-018-0326-x.
21. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 20 16;7:29. Published 2016 Jul 26: https://doi.org/10.4103/2153-3539.186902.
22. Retson TA, Besser AH, Sall S, Golden D, Hsiao A. Machine learning and deep neural networks in thoracic and cardiovascular imaging. j thorac imaging 2019;34(3):192–201. https://doi.org/10.1097/RTI.0000000000000385.
23. Chang H, Han J, Zhong C, Snijders AM, Mao JH. Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications. IEEE Trans Pattern Anal Mach Intell 2018;40(5):1182–1194. https://doi.org/10.1109/TPAMI.2017.2656884.
24. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018;9(4):611–629. https://doi.org/10.1007/s13244-018-0639-9.
25. Chang K, Bai HX, Zhou H, Su C, Bi WL. Agbodza E., et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. Clin Cancer Res 2018;24(5):1073–1081. https://doi.org/10.1158/1078-0432.CCR-17-2236.
26. Coudray N., Ocampo P.S., Sakellaropoulos T., Narula N., Snuderl M., Fenyö D., et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med. 2018;24(10):1559–1567. https://doi.org/10.1038/s41591-018-0177-5 (Lung).
27. Tabata K., Uraoka N., Benhamida J., Hanna M.G., Sirintrapun S.J., Gallas B.D., et al. Validation of mitotic cell quantification via microscopy and multiple whole-slide scanners. Diagn Pathol. 2019;14(1):65. Published 2019 Jun 26 https://doi.org/10.1186/s13000-019-0839-8.
28. Melo RCN, Raas MWD, Palazzi C, Neves VH, Malta KK, Silva TP. Whole slide imaging and its applications to histopathological studies of liver disorders. Front Med (Lausanne) 2020;6:310. Published 2020 Jan 8: https://doi.org/10.3389/fmed.2019.00310.
29. Zarella MD, Bowman D, Aeffner F, Farahani N, Xthona A. Absar S.F., et al. A practical guide to whole slide imaging: a white paper from the digital pathology association. Arch Pathol Lab Med 2019;143(2):222–234. https://doi.org/10.5858/arpa.2018-0343-RA.
30. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT. Way G.P., et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018;15(141):20170387. https://doi.org/10.1098/rsif.2017.0387.
31. Serag A, Ion-Margineanu A, Qureshi H, McMilian R, Martin M-JS. Diamon d J., et al. Translational AI and deep learning in diagnostic pathology. Front Med (Lausanne). 2019;6:185. Published 2019 Oct;1https://doi.org/10.3389/fmed.2019.00185.
32. Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676–682. Published 2012 Jun 28 https://doi.org/10.1038/nmeth.2019.
33. Nativ NI, Chen AI, Yarmush G, Henry SD, Lefkowtich J.H.,. Klein K.M., et al. Automated image analysis method for detecting and quantifying macrovesicular steatosis in hematoxylin and eosin-stained histology images of human livers. Liver Transpl 2014;20(2):228–236. https://doi.org/10.1002/lt.23782.
34. Girolami I., Parwani A., Barresi V., Marletta S., Ammendola S., Stefanizzi L., et al. The landscape of digital pathology in transplantation: from the beginning to the virtual E-slide. J Pathol Inform. 2019;10:21. Published 2019 Jul 1 https://doi.org/10.4103/jpi.jpi_27_19.
35. Edmondson HA, Steiner PE. Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 1954;7(3):462–503. https://doi.org/10.1002/1097-0142(195405)7:3<462::aid-cncr2820070308>3.0.co;2-e.

36. Atupelage C, Nagahashi H, Kimura F, Yamaguchi M, Tokiya A. Hashiguchi A., et al. Computational hepatocellular carcinoma tumor grading based on cell nuclei classification. J Med Imaging (Bellingham) 2014;1(3, 034501. https://doi.org/10.1117/1.JMI.1.3.034501.
37. Kiani A., Uyumazturk B., Rajpurkar P., Wang A., Gao R., Jones E., et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ Digit Med. 2020;3:23. Published 2020 Feb 26 https://doi.org/10.1038/s41746-020-0232-8.
38. Liao H., Long Y., Han R., Wang W., Xu L., Liao M., et al. Deep learning-based classification and mutation prediction from histopathological images of hepatocellular carcinoma [published online ahead of print, 2020 Jun 14]. Clin Transl Med. 2020. https://doi.org/10.1002/ctm2.102.
39. Schwen LO, Homeyer A, Schwier M, Dahmen U, Dirsch O, Schenk A, et al. Zonated quantification of steatosis in an entire mouse liver. Comput Biol Med 2016;73:108–118. https://doi.org/10.1016/j.compbiomed.2016.04.004.
40. Homeyer A., Hammad S., Schwen L.O., Dahmen U., Höfener H., Gao Y., et al. Focused scores enable reliable discrimination of small differences in steatosis. Diagn Pathol. 2018;13(1):76. Published 2018 Sep 20 https://doi.org/10.1186/s13000-018-0753-5.
41. Jedrzkiewicz J, Bronner MP, Salama ME, Kohan J, Rowe LR, Tai D, et al. Liver fibrosis quantification by digital whole slide imaging and two photon microscopy with second harmonic generation. Int J Pathol Clin Res 2018;4:078. https://doi.org/10.23937/2469-5807/1510078.
42. Cabibi D, Bronte F, Porcasi R, Ingrao S, Giannone AG, Maida M, et al. Comparison of histochemical stainings in evaluation of liver fibrosis and correlation with transient elastography in chronic hepatitis. Anal Cell Pathol (Amst) 2015;2015:431750. https://doi.org/10.1155/2015/431750.
43. Masugi Y., Abe T., Tsujikawa H., Effendi K., Hashiguchi A., Abe M., et al. Quantitative assessment of liver fibrosis reveals a nonlinear association with fibrosis stage in nonalcoholic fatty liver disease. Hepatol Commun. 2017;2(1):58–68. Published 2017 Nov 27 https://doi.org/10.1002/hep4.1121.
44. Barry B, Buch K, Soto JA, Jara H, Nakhmani A, Anderson SW. Quantifying liver fibrosis through the application of texture analysis to diffusion weighted imaging. Magn Reson Imaging 2014;32(1):84–90. https://doi.org/10.1016/j.mri.2013.04.006.
45. Simon O, Yacoub R, Jain S, Tomaszewski JE, Sarder P. Multi- radial LBP features as a tool for rapid glomerular detection and assessment in whole slide histopathology images. Sci Rep 2018;8(1):2032. Published 2018 Feb 1: https://doi.org/10.1038/s41598-018-20453-7.
46. Rosenberg A.Z., Palmer M., Merlino L., Troost J.P., Gasim A., Bagnasco S., et al. The application of digital pathology to improve accuracy in glomerular enumeration in renal biopsies. PLoS One. 2016;11(6):e0156441. Published 2016 Jun 16 https://doi.org/10.1371/journal.pone.0156441.
47. Barisoni L, Nast CC, Jennette JC, Hodgin JB, Herzenberg AM, Lemley KV, et al. Digital pathology evaluation in the multicenter Nephrotic Syndrome Study Network (NEPTUNE). Clin J Am Soc Nephrol 2013;8(8):1449–1459. https://doi.org/10.2215/CJN.08370812.
48. Tian K., Rubadue C.A., Lin D.I., Veta M., Pyle M.E., Irshad H., et al. Automated clear cell renal carcinoma grade classification with prognostic significance. PLoS One. 2019;14(10):e0222641. Published 2019 Oct 3 https://doi.org/10.1371/journal.pone.0222641.
49. Delahunt B. Advances and controversies in grading and staging of renal cell carcinoma. Mod Pathol 2009;22(suppl 2):S24–S36. https://doi.org/10.1038/modpathol.2008.183.
50. Yeh FC, Parwani AV, Pantanowitz L, Ho C. Automated grading of renal cell carcinoma using whole slide imaging. J Pathol Inform 2014;5(1):23. Published 2014 Jul 30: https://doi.org/10.4103/2153-3539.137726.
51. Furness PN, Levesley J, Luo Z, Taub N, Kazi JI, Bates WD, et al. A neural network approach to the biopsy diagnosis of early acute renal transplant rejection. Histopathology 1999;35(5):461–467. https://doi.org/10.1046/j.1365-2559.1999.035005461.
52. Wang S., Chen A., Yang L., Cai L., Xie Y., Fujimoto J., et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. Sci Rep. 2018;8(1):10393. Published 2018 Jul 10 https://doi.org/10.1038/s41598-018-27707-4.
53. Wang S., Yang D.M., Rong R., Zhan X., Fujimoto J., Liu H., et al. Artificial intelligence in lung cancer pathology image analysis. Cancers (Basel). 2019;11(11):1673. Published 2019 Oct 28 https://doi.org/10.3390/cancers11111673.
54. Yu K.H., Zhang C., Berry G.J., Altman R.B., Ré C., Rubin D.L., et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016;7:12474. Published 2016 Aug 16 https://doi.org/10.1038/ncomms12474.
55. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 2002;8(8):816–824. https://doi.org/10.1038/nm733.
56. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for cox's proportional hazards model via coordinate descent. J Stat Softw 2011;39(5):1-13. https://doi.org/10.18637/jss.v039.i05.
57. 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 Comput Soc Conf Comput Vis Pattern Recognit 2016;2016:2424–2433. https://doi.org/10.1109/CVPR.2016.266.
58. Teramoto A, Tsukamoto T, Kiriyama Y, Fujita H. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. Biomed Res Int 2017;2017:4067832. https://doi.org/10.1155/2017/4067832.
59. Pedraza A, et al. In: *Hernández M Valdés, González-Castro V, eds.* Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings. Springer International Publishing; 2017. p. 839–849.
60. Yu KH, Wang F, Berry GJ, Ré C, Altman RB, Snyder M, et al. Classifying non-small cell lung cancer types and transcriptomic subtypes using convolutional neural networks. J Am Med Inform Assoc 2020;27(5):757–769. https://doi.org/10.1093/jamia/ocz230.
61. Multi-Field-of-View Deep Learning Model Predicts Nonsmall Cell Lung Cancer Programmed Death-Ligand 1 Status from Whole-Slide Hematoxylin and Eosin Images.
62. Hanna N., Johnson D., Temin S., Baker S., Brahmer J., Ellis P.M., et al. Systemic therapy for stage IV non-small-cell lung cancer: American Society of Clinical Oncology Clinical Practice Guideline update [published correction appears in J Clin Oncol. 2018 Jan 20;36(3):304]. J Clin Oncol. 2017;35(30):3484–3515. https://doi.org/10.1200/JCO.2017.74.6065.
63. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep 2018;23(1):181-193.e7. https://doi.org/10.1016/j.celrep.2018.03.086.
64. Ichikawa T, Aokage K, Sugano M, Miyoshi T, Kojima M, Fujii S, et al. The ratio of cancer cells to stroma within the invasive area is a histologic prognostic parameter of lung adenocarcinoma. Lung Cancer 2018;118:30–35. https://doi.org/10.1016/j.lungcan.2018.01.023.
65. Yi F., Yang L., Wang S., Guo L., Huang C., Xie Y., et al. Microvessel prediction in H&E Stained Pathology Images using fully convolutional neural networks. BMC Bioinformatics. 2018;19(1):64. Published 2018 Feb 27 https://doi.org/10.1186/s12859-018-2055-z.
66. Wang S, Wang T, Yang L, Yang DM, Fujimoto J, Yi F, et al. ConvPath: a software tool for lung adenocarcinoma digital pathological image analysis aided by a convolutional neural network. EBioMedicine 2019;50:103–110. https://doi.org/10.1016/j.ebiom.2019.10.033.
67. Pantanowitz L, Sinard JH, Henricks WH, Fatheree FA, Carter AB, Contis L, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med 2013;137(12):1710–1722. https://doi.org/10.5858/arpa.2013-0093-CP.
68. Kim J, Bae JS. Tumor-associated macrophages and neutrophils in tumor microenvironment. Mediators Inflamm 2016;2016:6058147. https://doi.org/10.1155/2016/6058147.
69. Stanton SE, Disis ML. Clinical significan ce of tumor-infiltrating lymphocytes in breast cancer. J Immunother Cancer 2016;4:59. Published 2016 Oct 18: https://doi.org/10.1186/s40425-016-0165-6.
70. Moulin P, Grünberg K, Barale-Thomas E, der Laak JV. IMI-bigpicture: a central repository for digital pathology. Toxicol Pathol 2021 Jun;49(4):711–713. https://doi.org/10.1177/0192623321989644.
71. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis. Review Lancet Digit Health 2021 Mar;3(3):e195–e203. https://doi.org/10.1016/S2589-7500(20)30292-2.
72. Zhu S, Gilbert M, Chetty I, Siddiqui F. The 2021 landscape of FDA-approved artificial intelligen ce/machine learning-enabled medical devices: an analysis of the characteristics and intended use. Int J Med Inform 2022 Sep;165, 104828. https://doi.org/10.1016/j.ijmedinf.2022.104828. Epub 2022 Jul 1.
73. Epub 2021 Mar 8Farris AB, Vizcarra J, Amgad M, Cooper LAD, Gutman D, Hogan J. Artificial intelligence and algorithmic computational pathology: an introduction with renal allograft examples. Histopathology 2021 May;7 8(6):791–804. https://doi.org/10.1111/his.14304.
74. Farris AB, Vizcarra J, Amgad M, Donald Cooper LA, Gutman D, Hogan J. Image analysis pipeline for renal allograft evaluation and fibrosis quantification. Kidney Int Rep 2021 Apr 24;6(7):1878–1887. https://doi.org/10.1016/j.ekir.2021.04.019.
75. Hermsen M., Volk V., Bräsen J.H., Geijs D.J., Gwinner W., Kers J., et al. Quantitative assessment of inflammatory infiltrates in kidney transplant biopsies using multiplex tyramide signal amplification and deep learning. Lab Invest. 2021 Aug;101(8):970–982. Epub 2021 May 18. https://doi.org/10.1038/s41374-021-00601-w.
76. Brunt E.M., Clouston A.D., Goodman Z., Guy C., Kleiner D.E., Lackner C., et al. Complexity of ballooned hepatocyte feature recognition: defining a training atlas for artificial intelligence-based imaging in NAFLD. J Hepatol. 2022 May;76(5):1030–1041. Epub 2022 Jan 25. https://doi.org/10.1016/j.jhep.2022.01.011.
77. Naoumov N.V., Brees D., Loeffler J., Chng E., Ren Y., Lopez P., et al. Digital pathology with artificial intelligence analyses provides greater insights into treatment-induced fibrosis regression in NASH. J Hepatol. 2022 Nov;77(5):1399–1409. Epub 2022 Jun 30. https://doi.org/10.1016/j.jhep.2022.06.018.
78. Cherchi V., Mea V.D., Terrosu G., Brollo P.P., Pravisani R., Calandra S., et al. Assessment of hepatic steatosis based on needle biopsy images from deceased donor livers. Clin Transplant. 2022 Mar;36(3):e14557. Epub 2022 Jan 5. https://doi.org/10.1111/ctr.14557.
79. Roy M., Wang F., Vo H., Teng D., Teodoro G., Farris A.B., et al. Deep-learning-based accurate hepatic steatosis quantification for histological assessment of liver biopsies. Lab Invest. 2020 Oct;100(10):1367–1383. Epub 2020 Jul 13. https://doi.org/10.1038/s41374-020-0463-y.
80. Yu H., Sharifai N., Jiang K., Wang F., Teodoro G., Farris A.B., et al. Artificial intelligence based liver portal tract region identification and quantification with transplant biopsy whole-slide images. Comput Biol Med. 2022 Sep 6;150:106089. Epub ahead of print. https://doi.org/10.1016/j.compbiomed.2022.106089.
81. Yang T.L., Tsai H.W., Huang W.C., Lin J.C., Liao J.B., Chow N.H., et al. Pathologic liver tumor detection using feature aligned multi-scale convolutional network. Artif Intell Med. 2022 Mar;125:102244. Epub 2022 Jan 10. https://doi.org/10.1016/j.artmed.2022.102244.
82. Uegami W., Bychkov A., Ozasa M., Uehara K., Kataoka K., Johkoh T., et al. MIXTURE of human expertise and deep learning-developing an explainable model for predicting pathological diagnosis and survival in patients with interstitial lung disease. Mod Pathol. 2022 Aug;35(8):1083–1091. Epub 2022 Feb 23. https://doi.org/10.1038/s41379-022-01025-7.
83. Qaiser T., Lee C.Y., Vandenberghe M., Yeh J., Gavrielides M.A., Hipp J., et al. Usability of deep learning and H&E images predict disease outcome-emerging tool to optimize clinical trials. NPJ Precis Oncol. 2022 Jun 15;6(1):37. https://doi.org/10.1038/s41698-022-00275-7.
84. Rączkowski Ł., Paśnik I., Kukiełka M., Nicos M., Budzinska M.A., Kucharczyk T., et al. Deep learning-based tumor microenvironment segmentation is predictive of tumor mutations and patient survival in non-small-cell lung cancer. BMC Cancer.. 2022 Sep 21;22(1):1001. https://doi.org/10.1186/s12885-022-10081-w.