



Whole Slide Imaging: Deep Learning and Artificial Intelligence

Asmaa Aljuhani, Arunima Srivastava, James P. Cronin, Jany Chan, Raghu Machiraju, and Anil V. Parwani

Abbreviations

AI	Artificial intelligence
CNN	Convolutional neural networks
DL	Deep learning
DNN	Deep convolutional neural networks
HSV	Hue, saturation, value
ML	Machine learning
RGB	Red, green, blue

A. Aljuhani (✉) · A. Srivastava · J. Chan
Department of Computer Science and Engineering, College of Engineering,
The Ohio State University, Columbus, OH, USA
e-mail: aljuhani.2@osu.edu; srivastava.92@osu.edu; chan.206@osu.edu

J. P. Cronin
Department of Veterinary Biosciences, College of Veterinary Medicine,
The Ohio State University, Columbus, OH, USA
Department of Internal Medicine, College of Medicine, The Ohio State University,
Columbus, OH, USA
e-mail: james.cronin@osumc.edu

R. Machiraju
Department of Computer Science and Engineering, College of Engineering,
The Ohio State University, Columbus, OH, USA
Department of Biomedical Informatics, College of Medicine, The Ohio State University,
Columbus, OH, USA

Department of Pathology, College of Medicine, The Ohio State University Wexner Medical
School, Columbus, OH, USA
e-mail: machiraju.1@osu.edu

A. V. Parwani
Department of Pathology, College of Medicine, The Ohio State University Wexner Medical
School, Columbus, OH, USA
e-mail: anil.parwani@osumc.edu

ROC	Receiver operating characteristic
ROI	Regions of interest
TIFF	Tagged Image File Format
WSI	Whole slide image

Introduction

Converging waves of increasingly sophisticated machine learning (ML), whole slide images (WSIs), and computing power have artificial intelligence (AI) (Fig. 1) poised to transform the practice of pathology. The ongoing collaboration of researchers from computer vision, AI, and pathology domains is driving this revolution. A recent explosion of ML models for the analysis of WSIs has produced state-of-the-art biomarker discoveries and impressive disease recognition capabilities [1]. ML has the potential to address the worsening global undersupply of pathologists [2] and the thorny issue of interpathologist variability [3]. Additionally, ML can be used to optimize the diagnostic pathologist's workflow via (1) attention direction to regions of interest (ROI) and (2) automated quantification of time-intensive tasks (e.g., mitotic indices). From the discovery perspective, ML can identify novel features of WSIs with prognostic and therapeutic significance in a variety of neoplastic and metabolic conditions [4, 5].

ML can be unsupervised or supervised. *Unsupervised* models do not introduce labeling bias when learning patterns in data. Rather, the model identifies distinct patterns in the data and forms clusters with unique patterns. Unsupervised learning is useful in an exploratory analysis in which ground truth is unknown. In comparison, *supervised* learning utilizes manually assigned labels from ground truth that identify relevant features of the dataset. Supervised models are conducive to iterative improvement, as the presence of labels helps optimize the model. The performance of the

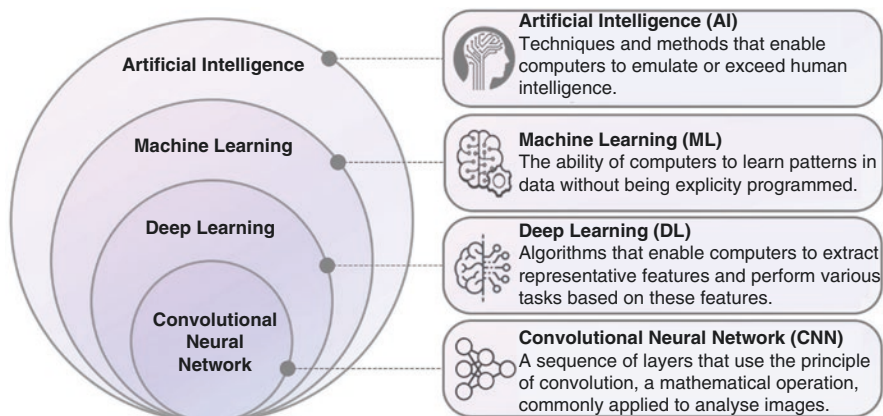


Fig. 1 The hierarchical relationship of different artificial intelligence concepts

supervised model depends on the (1) features, (2) labels, and (3) core algorithm used in training.

Deep learning (DL) is a subcategory of ML (Fig. 1) known for its ability to achieve high performance from complex visual inputs, such as WSIs [6]. DL algorithms utilize networks several layers in depth, progressively extracting higher level features from the raw input with each additional layer. DL algorithms iteratively improve by maximizing the separation between classes. With each iteration, data are propagated through the network to determine the corresponding output. The machine-predicted output is then compared to the actual output, and a penalty score is assigned so that the algorithm can learn to map the sample output to the correct class. Once the algorithm determines the discriminant features for each class, it is often able to generalize to unseen data without the need to handcraft additional features.

The convolutional neural network (CNN) is typically a *supervised* method under the DL umbrella (Fig. 1) that has recently been applied to digital pathology. CNNs are generally used to analyze images, where they assign weights to different regions and structures to model and classify groups. CNNs use the principle of convolution, in which a mathematical operation on two functions is used to produce a third that highlights essential structures (i.e., changes in signal or an underlying smoothness). CNNs are composed of three main types of layers: convolution layers, pooling layers, and fully connected layers. Stacking these layers forms a CNN architecture. The more layers added, the “deeper” the network becomes, hence the name deep convolutional neural networks (DNNs).

In this chapter, we highlight challenges in implementing CNNs in digital pathology (section “[Challenges in Implementing Convolutional Neural Networks in Digital Pathology](#)”), discuss data quality and transformation (section “[Data Quality and Transformation](#)”), inform annotation and labeling (section “[Annotation and Labeling](#)”), demystify CNNs (section “[Convolutional Neural Networks](#)”), explore fine-tuning CNNs (section “[Further Steps for Fine-Tuning the CNN](#)”), and list modern applications for AI in digital pathology (section “[Applications of AI in WSI](#)”).

Challenges in Implementing Convolutional Neural Networks in Digital Pathology

Computational modeling of WSIs poses many unique challenges. CNNs are data-driven and require large datasets for training, validating, and testing. The development of large, high-quality datasets is impeded by several barriers to entry in digital pathology, including cost, expertise, and resistance to change. There are multiple steps in data pre-processing with the goal of maintaining data quality and optimizing data transformation. A compatible image format is imperative for downstream analysis, and investigators should consider the entire pipeline before selecting the image format. Different scanners can use propriety data formats for both image generation and annotation, which can add unique challenges for pre-processing and analysis. Investigators can choose from a variety of color spaces, transformations,

and contrasts to suit their purpose. The images must then be tiled and filtered with care to maintain the representation of all structures of interest. Normalization is required to counteract batch effects, which can increase image variability due to disparate sample handling. The challenge of isolating distinct morphologic features can be overcome via stain deconvolution, a powerful computational technique for isolating the relative contributions of hematoxylin and eosin staining.

After image pre-processing, pathologist expertise is required to annotate key features for training. Pathology is a highly specialized field, and different organs and diseases require pathologists with a variety of specializations in order to generate accurate annotations. Furthermore, image annotation is time-consuming and requires multiple pathologists to reach a consensus [7]. For any computational modeling endeavor to be executed successfully in the histology domain, the modeling approach must be designed with the input of an expert pathologist at every stage. Hence, each modeling effort should begin with the well-understood integration of pathologists. Pathologist expertise to annotate data, construct models, and verify results is of utmost importance to ensure usability and adoption of AI in pathology.

Ultimately, careful consideration of the parameters for the modeling algorithms, the feature sets, and the neural network architecture are all essential pieces in the overall success of a digital pathology modeling experiment. From the size of the tiles (must contain enough of the relevant tissue substructures but not so much as to add unnecessary variation and noise) to the complexity of the ML model (less training data with more complex models leads to overfitting), all decisions impact the results and should be made after careful consideration and comprehensive validation [8].

Data Quality and Transformation

Sample Size

In computer-aided pathology, the size of the dataset is a crucial factor underlying model performance. The more data fed into the algorithm, the more accurately it will be able to model the full range of the disease of interest. Variation in the form of disease presentation and processing techniques must be captured in training to ensure robust results.

Image Format

Digital pathology relies on scanning hardware to convert glass slides into specific image formats with high resolutions. Automated image processors use existing standard formats or unique proprietary formats with associated tools and viewers [9]. Generally, the difference between formats stems from different metadata tags used, as well as the file compression type. Investigators should be aware that downstream analysis depends on how well computational tools handle the chosen image format. For example, fast rendering in the viewer, ease of annotations, and data

management are dependent on the file format. Converting from a scanner-specific format to a standard format may be possible. However, lossy compression methods that degrade the data may be required to achieve a smaller size that is capable of easy viewing.

A standard image format for WSIs is the TIFF (Tagged Image File Format) with lossless compression to maintain image details via storage as multi-resolution (or “pyramidal”) representations [10]. Scanner-specific formats include SVS (based on TIFF) from Aperio scanners [11] and MRXS from the Zeiss MIRAX series [12] and 3DHISTECH Panoramic series [13]. These files typically contain multiple images that range from full-resolution to a low-resolution thumbnail [14]. Any or all these images can be extracted, and the investigator’s choice will depend on the resolution needed for analysis.

Color Space, Transformation, and Contrast

Many downstream analyses, such as segmentation and object counting, are based on native *color space*. Thus, *transformation* of an image to a different color space affects the results of these endeavors. Different color spaces focus on distinct image quality characteristics. To illustrate, RGB (red, blue, green) and HSV (hue, saturation, value) are shown in Fig. 2. The number of possible color spaces is too vast to list here, and an investigator’s selection will be informed by their objective. A straightforward and commonly used transformation is color to grayscale. This transformation has one feature per pixel: color intensity. Standard ML enables edge detection and segmentation using color intensity and can facilitate precise homogeneous region identification [15]. Similarly, a change in the *contrast* of an image can enable the detection of larger, more apparent objects. A change in contrast essentially changes the difference in luminance between objects in the image. In a grayscale image, darker objects become darker and lighter objects become lighter, in some cases rendering subtle details more apparent [16].

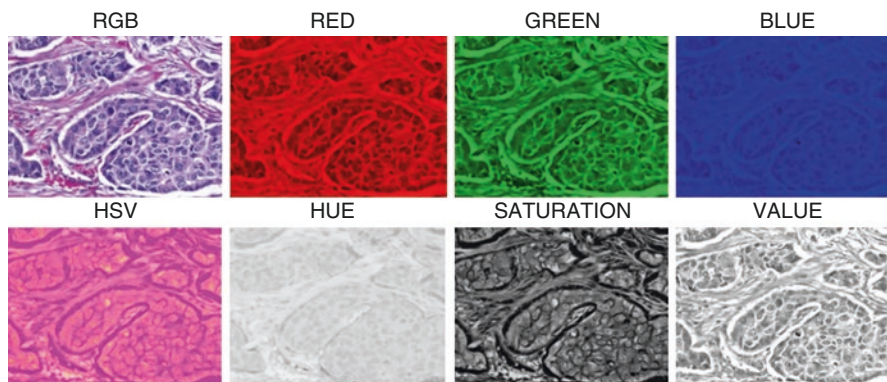


Fig. 2 RGB vs. HSV color spaces and their individual channels for a digital image of breast cancer

Tiling and Filtering the Image

In most cases, the whole image should be *tiled* for faster processing, meaning the whole image is segmented into smaller, rectangular regions or *tiles*, and irrelevant parts of the image should be *filtered*. The size of the tiles needs to be appropriate for the analysis being performed. Since the tile analyses are done in lieu of analyzing the entire WSI, the tiles need to be representative of the structures present in the whole tissue. Thus, the location, size, and magnification should facilitate each tile containing relevant structures [17].

Images can be *filtered* in multiple ways. We can filter artifacts (e.g., white space) or biological entities that do not pertain to the question (e.g., non-tumor region). The easiest way to perform filtration is to compute a measure per tile, which denotes whether the tile is useful or not. For example, if we aimed to analyze any areas which were not predominantly white space, we could average the RGB values of all pixels in each tile and use a threshold to demarcate the tiles to be included in the analysis. An alternative method is to extract specific ROIs from each tile and discard the rest of the image.

Normalization

A standard step in any data modeling protocol is data *normalization*, and computational modeling of WSIs is no different. Normalization is required whenever a set of images is to be analyzed together. This step is imperative as WSIs exhibit considerable variation and are highly prone to batch effects. Sources of variability include histology lab personnel, staining procedures, lab instruments, scanners, and digitization protocols [18]. Most normalization techniques transform all slides in the dataset to mimic a preselected reference slide [19]. The reference slide needs to be an accurate representation of the staining and structures across all slides. Hence, choosing a reference slide poses a challenge. Normalization techniques include pixel-wise standardization of image colors, brightness, and contrast. There are multiple proposed computational methods to perform normalization, and newer, more sophisticated methods are being developed using neural networks [20]. Approaches include color space transformation in the RGB space and color deconvolution that isolates the contribution of the two stain vectors, hematoxylin and eosin.

Stain Deconvolution

Since hematoxylin and eosin dyes adhere to different tissue components, an important step of many analysis protocols is to separate these two dyes in the image. This results in two grayscale images, one of each stain (Fig. 3). For some downstream analyses, such as counting nuclei, distinguishing epithelium and stroma, and assessing the nuclear to cytoplasmic ratio, single-channel grayscale is a powerful technique.

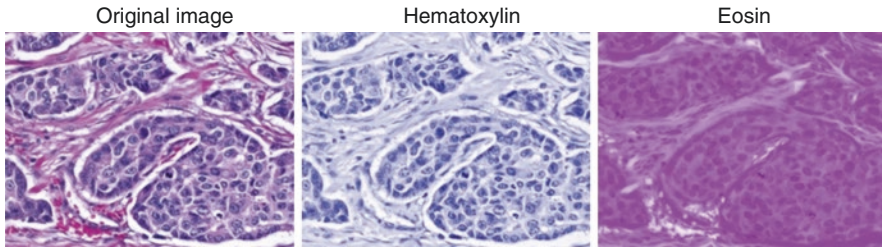


Fig. 3 An example of applying color deconvolution on a digital image of breast cancer

There are a variety of methods for stain deconvolution [21]. Most use a *stain matrix*, which, when multiplied with the color space channels, will produce a stain channel. These channels are specific to each image (or a set of images if they are normalized) and can be transformed into a grayscale image that represents stain intensity.

Annotation and Labeling

Annotation

Pathologists evaluate various structural, textural, and morphological markers to find evidence of disease. This expertise is achieved through years of training. Similarly, CNNs must be trained to identify diagnostic features and to ignore irrelevant noise and artifacts. This is accomplished via *annotation* of key morphological features. The specific features that are labeled depend on the problem to be addressed. For example, annotation of mitotic cells can inform a model predicting tumor grade in breast cancer [22].

Ideally, annotation protocols are determined at the inception of the computational modeling project with consideration of the clinical question/problem. Depending on the task, various implementations may be suitable, such as point annotations (that identify the centroid of the pathology marker), shape annotations (that define a bounding pre-defined shape around the pathology marker), or granular outline annotations (that precisely segment out the pathology marker). A categorical label needs to be assigned to each annotation. An annotation tool that allows for viewing the WSI, efficient annotation, and exportation is required. Annotation tools and software are commercially available with some provided by the image scanner manufacturers [23]. There are several open-source tools that support WSIs in a variety of formats, including QuPath, HistomicsTK, and ASAP. The annotations are exported in easily interpretable text formats such as JSON and XML [24]. Some annotation tools provide options for automated analysis, image normalization, and segmentation to aid in more efficient annotation of many images.

Recent crowdsourcing initiatives for histology annotation have been successful in aggregating labels for large datasets [25]. This has helped to address limitations in dataset sample size. Such initiatives facilitate computational modeling solutions that could be benchmarked and repurposed by computational pathologists to better understand and model in-house data [22].

Clinical and Histopathologic Labels

In contrast to annotation, which assigns labels to specific morphologic features, each WSI may be assigned a diagnostic label for training. Labels are shared by all the tiles emanating from the WSI. Examples of diagnostic labels include disease subtype, grade, sequencing data, drugs administered, and survival. These labels are extracted from patient records or derived by a subject matter expert who reviews the data before processing. Using labeling for sequencing data, we can identify recurring patterns that characterize genetic subtypes of the disease [26].

Convolutional Neural Networks

Prior to DL, traditional classification approaches required researchers to manually harvest domain-specific features. This process of extracting handcrafted features required extensive tuning to accommodate the variability of the data, and applicability to other problems (i.e., analyzing different diseases) was limited. Addressing this challenge, DL follows a *domain agnostic* approach, combining the process of automated feature extraction with the identification of discriminating markers. Thus, the process of harvesting discriminatory features becomes automated.

Deep convolutional neural networks (DNNs) (Fig. 4) have a dominant learning ability due to multiple feature extraction stages that allow them to learn representative features of the data. This powerful capability has earned DNNs steady popularity in analyzing large, high-resolution WSIs across a variety of cancer subtypes [27], as well as many other conditions, such as Alzheimer's disease [28].

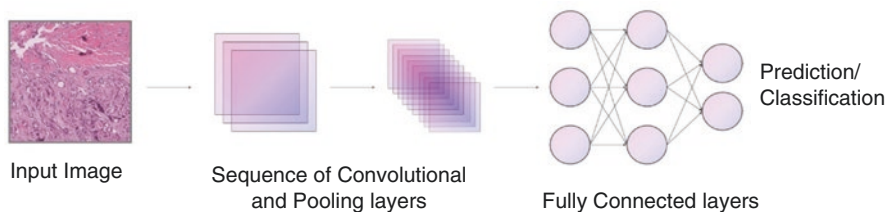


Fig. 4 The common structure of DNN models. An image is passed through a series of convolutional and pooling layers. These layers extract representative features that are used in the fully connected layers to classify the input image

Anatomy of a CNN

Neurons, the basic building block of the neural network, are assigned to one of three possible layers: *input*, *hidden*, or *output*. If every input neuron is connected to every output neuron and vice versa, the layers are considered **fully connected**. The input layer receives a pre-processed image as a matrix and passes it to the first hidden layer. The hidden layers perform mathematical computations to extract relevant image features. Lastly, the output layer returns the predicted value for the input image based on the features identified in the hidden layers.

Each connection between neurons is associated with a **weight** that prescribes the importance of the value from a neuron in the preceding layer. These are called the model's **parameters**. At the beginning of the training process, these weights are assigned randomly. Throughout the learning process, the model adjusts these weights based on how accurately it predicts the actual outputs. A loss function is used to evaluate the learning ability of the model. Ideally, the generated loss function is close to zero, which means the labels generated by the model are highly correlated with the actual labels.

Hidden Layers

The convolutional layer is the first layer to extract features from an input image. It preserves the dimensions of the input. It is based on a mathematical operation that takes two inputs, such as an image and a filter, and produces a convolved feature output. Applying different types of filters can generate the following transformations: edge detection, smoothing, and sharpening the input image.

The pooling layer is used to reduce the dimensionality of the input image to shorten training time and combat overfitting. There are different types of special pooling: max pooling, average pooling, and sum pooling.

The activation layer operates to minimize a loss function. This layer is to classify the output into different classes. The choice of activation function is dependent on the desired output. For example, sigmoid is preferred for binary classification while softmax is typically used for multiple classes [29].

Further Steps for Fine-Tuning the CNN

Feature Identification

A feature is defined as any measurable property of the WSI that is characteristic of the phenomenon being observed. For example, features can define a cell nucleus, inflammatory cells, extracellular matrix, etc. Choosing discriminative, informative, and independent features is a crucial component of developing a powerful CNN. The inherent statistics of the feature set, such as variation and distance between data points in the feature space (e.g., Euclidean space), will be used by the CNN to

predict the most appropriate label (*supervised* models) or grouping (*unsupervised* models) for individual data points.

Validation and Performance

For model *validation*, common measurements like accuracy, precision, recall, F-score, and mean squared error evaluate correctness in different contexts. To assess the *performance* of a model, one can utilize K-fold validation, randomization of the input data, or titration with noise to compare the penultimate results. Techniques such as the receiver operating characteristic (ROC) curves make these measures and the changes easy to interpret and contextualize.

Applications of AI in WSI

Detection and Segmentation

CNNs facilitate the *detection* of disease-relevant structures and the subsequent *segmentation* of ROIs with high probability. This capability allows CNNs to be used as pre-screening and augmentation tools during histopathologic diagnosis of digitized slides. The CNN-guided discovery focuses the pathologist's attention on ROIs, thereby optimizing the pathologist's workflow. Moreover, identification and quantification of disease markers become standardized, thus reducing interpathologist variability.

The nucleus has been the target of many early studies in CNN segmentation. Investigators have successfully used several unique approaches and architectures to identify nuclear ROIs. For example, a PMap approach using CNNs gauges the probability of each pixel's proximity, according to its intensity, to cell nuclei to determine nuclei locations [30]. Alternatively, Mask R-CNN utilizes a region proposal network, first zeroing in on the areas that may contain nuclei and iteratively finding their exact boundaries for nuclei detection [31].

In addition to cellular features, detection of unique cellular phenomena, such as mitoses, is enabled by CNN segmentation. A standard method for quantifying mitotic figures is the mitotic count. Counting mitoses requires the pathologist to (1) identify the tumor region with highest mitotic activity, (2) differentiate mitoses from nuclear pyknosis, and (3) count mitotic events in at least ten representative, non-overlapping high-power fields. Each of these challenges is both time-consuming and highly prone to interobserver variability. DL networks that use spatial context to identify mitosis using a max pooling CNN have achieved significant success in mitosis identification [32]. A CNN feature set combined with domain-specific hand-crafted features gave rise to a computationally economical model which successfully identified mitosis [27].

CNNs have also been used to identify broad areas that contain multiple ROIs. For instance, to recognize tissue alternations of nonalcoholic fatty liver disease, a CNN model attained almost 95% accuracy, paving the way for more feasible and

rapid diagnosis [33]. A CNN trained on patch annotations to identify ROIs and post-process the segmentation with a fully connected conditional random field can be used to build a generalizable method for identifying regions of diagnostic relevance in histology images [34].

Artifact Discovery

There are a variety of *artifacts* in histology slides that can impede accurate computational diagnosis and hamper experts when using digitized WSIs. To address this challenge, CNNs can be used as quality control and correction tools. For example, a tool trained on different amounts of blurry histology and immunohistochemistry images can reliably identify artifactual ROIs. Tools such as these may soon be integrated with scanners to automatically re-scan artifactual ROIs and optimize the preparation process prior to pathologist interpretation [35].

Classification (Diagnosis and Grading)

Diagnosis and grading are *classification* tasks conducted by the pathologist in daily practice. There are many examples of CNNs achieving success in this domain, showcasing the potential to reduce interpathologist discordance and hasten accurate diagnosis. For example, a CNN-based method presented an accuracy of 98% when using confidence-based scoring from a deep network to classify histology tissue of skin cancer into four main classes [36]. Another model, performing predictions on patches of WSIs using CNNs and subsequently aggregating these, was able to deliver whole slide classification close to pathologist decisions for subtypes of cancers [37]. Extending the classification paradigm, tumor grades can be identified using CNNs, as is evidenced by studies in the kidney [38], brain [39], and prostate [40]. A CNN training framework with the relevant labeled images can go a step further and directly prognosticate using WSIs from cancer [41].

Summary

This is an exciting time in pathology diagnostics. CNNs are powerful tools for complex image analysis, making them ideal for digital pathology applications. The workflow of CNN development on WSIs has several challenges, and the perspective of the pathologist is welcome at every stage of model development. As it is widely deployed and adopted in clinical settings, WSI technology will allow pathologists to rapidly access and share images easily. Once well integrated with clinical workflows, WSI will be increasingly used in CNNs and AI applications for feature selection, tumor diagnosis, tumor grading, and developing image-based prognostic assays. Progress in CNN- and AI-based tool development will be further accelerated as overall WSI adoption for primary diagnosis and other clinical applications moves forward.

References

1. Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: a survey. arXiv:191212378 [cs, eess] [Internet]. 2019 Dec 27 [cited 2020 Jul 27]. Available from: <http://arxiv.org/abs/1912.12378>.
2. Robboy SJ, Weintraub S, Horvath AE, Jensen BW, Alexander CB, Fody EP, et al. Pathologist workforce in the United States: I. development of a predictive model to examine factors influencing supply. *Arch Pathol Lab Med*. 2013;137(12):1723–32.
3. Dasari S, Chakraborty A, Truong L, Mohan C. A systematic review of interpathologist agreement in histologic classification of lupus nephritis. *Kidney Int Rep*. 2019;4(10):1420–5.
4. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep*. 2018;8(1):3395.
5. Signaevsky M, Prastawa M, Farrell K, Tabish N, Baldwin E, Han N, et al. Artificial intelligence in neuropathology: deep learning-based assessment of tauopathy. *Lab Invest*. 2019;99(7):1019–29.
6. Lakhani P, Gray DL, Pett CR, Nagy P, Shih G. Hello world deep learning in medical imaging. *J Digit Imaging*. 2018;31(3):283–9.
7. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng*. 2009;2:147–71.
8. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J*. 2018;16:34–42.
9. Zarella MD, Bowman D, Aeffner F, Farahani N, Xthona A, Absar SF, et al. A practical guide to whole slide imaging: a white paper from the digital pathology association. *Arch Pathol Lab Med*. 2019;143(2):222–34.
10. Besson S, Leigh R, Linkert M, Allan C, Burel J-M, Carroll M, et al. Bringing open data to whole slide imaging. In: Reyes-Aldasoro CC, Janowczyk A, Veta M, Bankhead P, Sirinukunwattana K, editors. *Digital pathology* [Internet]. Cham: Springer International Publishing; 2019 [cited 2020 Jul 27]. p. 3–10. (Lecture Notes in Computer Science; vol. 11435). Available from: http://link.springer.com/10.1007/978-3-030-23937-4_1.
11. Grove LBD of LMI 1700 LLB, Fax: +1 847-236-3009 I 60089 USOP +1 844 534 2262. Scan – Aperio digital pathology slide scanners [Internet]. Leica Biosystems. [cited 2020 Jul 27]. Available from: <https://www.leicabiosystems.com/digital-pathology/scan/>.
12. Carl Zeiss: MIRAX LIVE [Internet]. Microscopy news. 2017 [cited 2020 Jul 27]. Available from: <https://microscopy-news.com/products/systems/carl-zeiss-mirax-live/>.
13. Digital pathology | EpreDia [Internet]. [cited 2020 Jul 27]. Available from: <https://epredia.com/digital-pathology-solutions/>.
14. Aperio format [Internet]. [cited 2020 Jul 27]. Available from: <https://openslide.org/formats/aperio/>.
15. Canny J. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell*. 1986;PAMI-8(6):679–98.
16. Ibrahim H, Pik Kong NS. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Trans Consum Electron*. 2007;53(4):1752–8.
17. Gertych A, Swiderska-Chadaj Z, Ma Z, Ing N, Markiewicz T, Cierniak S, et al. Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Sci Rep* [Internet]. 2019 Feb 6 [cited 2020 Jul 27];9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6365499/>.
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. Zanjani FG, Zinger S, Bejnordi BE, van der Laak JAWM, de With PHN. Stain normalization of histopathology images using generative adversarial networks. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018. p. 573–7.

20. Pontalba JT, Gwynne-Timothy T, David E, Jakate K, Androustos D, Khademi A. Assessing the impact of color normalization in convolutional neural network-based nuclei segmentation frameworks. *Front Bioeng Biotechnol.* 2019;7:300.
21. Alsubaie N, Trahearn N, Raza SEA, Snead D, Rajpoot NM. Stain deconvolution using statistical analysis of multi-resolution stain colour representation. *PLoS One.* 2017;12(1):e0169875.
22. Chang JM, McCullough AE, Dueck AC, Kosiorek HE, Ocal IT, Lidner TK, et al. Back to basics: traditional Nottingham grade mitotic counts alone are significant in predicting survival in invasive breast carcinoma. *Ann Surg Oncol.* 2015;22 Suppl 3:S509–15.
23. Grove LBD of LMI 1700 LLB, Fax: +1 847-236-3009 I 60089 USOP +1 844 534 2262. Aperio ImageScope - pathology slide viewing software [Internet]. Leica Biosystems. [cited 2020 Jul 27]. Available from: <https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/>.
24. ASAP - Automated slide analysis platform [Internet]. [cited 2020 Jul 27]. Available from: <https://computationalpathologygroup.github.io/ASAP/>.
25. Grote A, Schaadt NS, Forestier G, Wemmert C, Feuerhake F. Crowdsourcing of histological image labeling and object delineation by medical students. *IEEE Trans Med Imaging.* 2019;38(5):1284–94.
26. Coudray N, Ocampo PS, 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):1559–67.
27. Wang H, Cruz-Roa A, Basavanthally A, Gilmore H, Shih N, Feldman M, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *J Med Imaging (Bellingham).* 2014;1(3):034003.
28. Tang Z, Chuang KV, DeCarli C, Jin L-W, Beckett L, Keiser MJ, et al. Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline. *Nat Commun.* 2019;10(1):2173.
29. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. *Advances in neural information processing systems 25* [Internet]. Curran Associates, Inc.; 2012 [cited 2020 Jul 27]. p. 1097–1105. Available from: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
30. Höfener H, Homeyer A, Weiss N, Molin J, Lundström CF, Hahn HK. Deep learning nuclei detection: a simple approach can deliver state-of-the-art results. *Comput Med Imaging Graph.* 2018;70:43–52.
31. Jung H, Lodhi B, Kang J. An automatic nuclei segmentation method based on deep convolutional neural networks for histopathology images. *BMC Biomed Eng.* 2019;1(1):24.
32. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv.* 2013;16(Pt 2):411–8.
33. Arjmand A, Angelis CT, Christou V, Tzallas AT, Tsiouras MG, Glavas E, et al. Training of deep convolutional neural networks to identify critical liver alterations in histopathology image samples. *Appl Sci.* 2020;10(1):42.
34. Chan L, Hosseini M, Rowsell C, Plataniotis K, Damaskinos S. HistoSegNet: semantic segmentation of histological tissue type in whole slide images. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV); 2019. p. 10661–70.
35. Senaras C, Niazi MKK, Lozanski G, Gurcan MN. DeepFocus: detection of out-of-focus regions in whole slide digital images using deep learning. *PLoS One.* 2018;13(10):e0205387.
36. Ianni JD, Soans RE, Sankarapandian S, Chamarthi RV, Ayyagari D, Olsen TG, et al. Tailored for real-world: a whole slide image classification system validated on uncurated multi-site data emulating the prospective pathology workload. *Sci Rep.* 2020;10(1):3217.
37. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 2424–33.

38. Khoshdeli M, Borowsky A, Parvin B. Deep learning models differentiate tumor grades from H E stained histology sections. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2018. p. 620–3.
39. Ertosun MG, Rubin DL. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. *AMIA Annu Symp Proc.* 2015;2015:1899–908.
40. Li W, Li J, Sarma KV, Ho KC, Shen S, Knudsen BS, et al. Path R-CNN for prostate cancer diagnosis and gleason grading of histological images. *IEEE Trans Med Imaging.* 2019;38(4):945–54.
41. Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis C-A, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med.* 2019;16(1):e1002730.