

Review

# Objective Diagnosis for Histopathological Images Based on Machine Learning Techniques: Classical Approaches and New Trends

Naira Elazab <sup>1</sup>, Hassan Soliman <sup>1</sup>, Shaker El-Sappagh <sup>2,3</sup>, S. M. Riazul Islam <sup>4,\*</sup> and Mohammed Elmogy <sup>1</sup>

<sup>1</sup> Information Technology Department, Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt; naira.elazab@mans.edu.eg (N.E.); hsoliman@mans.edu.eg (H.S.); melmogy@mans.edu.eg (M.E.)

<sup>2</sup> Centro Singular de Investigación en Tecnologías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, 15705 Santiago de Compostela, Spain; shaker.elsappagh@usc.es

<sup>3</sup> Information Systems Department, Faculty of Computers and Artificial Intelligence, Benha University, Benha 13512, Egypt

<sup>4</sup> Department of Computer Science and Engineering, Sejong University, Seoul 05006, Korea

\* Correspondence: riaz@sejong.ac.kr; Tel.: +82-02-3408-2969

Received: 21 September 2020; Accepted: 19 October 2020; Published: 24 October 2020



**Abstract:** Histopathology refers to the examination by a pathologist of biopsy samples. Histopathology images are captured by a microscope to locate, examine, and classify many diseases, such as different cancer types. They provide a detailed view of different types of diseases and their tissue status. These images are an essential resource with which to define biological compositions or analyze cell and tissue structures. This imaging modality is very important for diagnostic applications. The analysis of histopathology images is a prolific and relevant research area supporting disease diagnosis. In this paper, the challenges of histopathology image analysis are evaluated. An extensive review of conventional and deep learning techniques which have been applied in histological image analyses is presented. This review summarizes many current datasets and highlights important challenges and constraints with recent deep learning techniques, alongside possible future research avenues. Despite the progress made in this research area so far, it is still a significant area of open research because of the variety of imaging techniques and disease-specific characteristics.

**Keywords:** medical image analysis; histopathology image analysis; conventional machine learning methods; deep learning methods; computer-assisted diagnosis

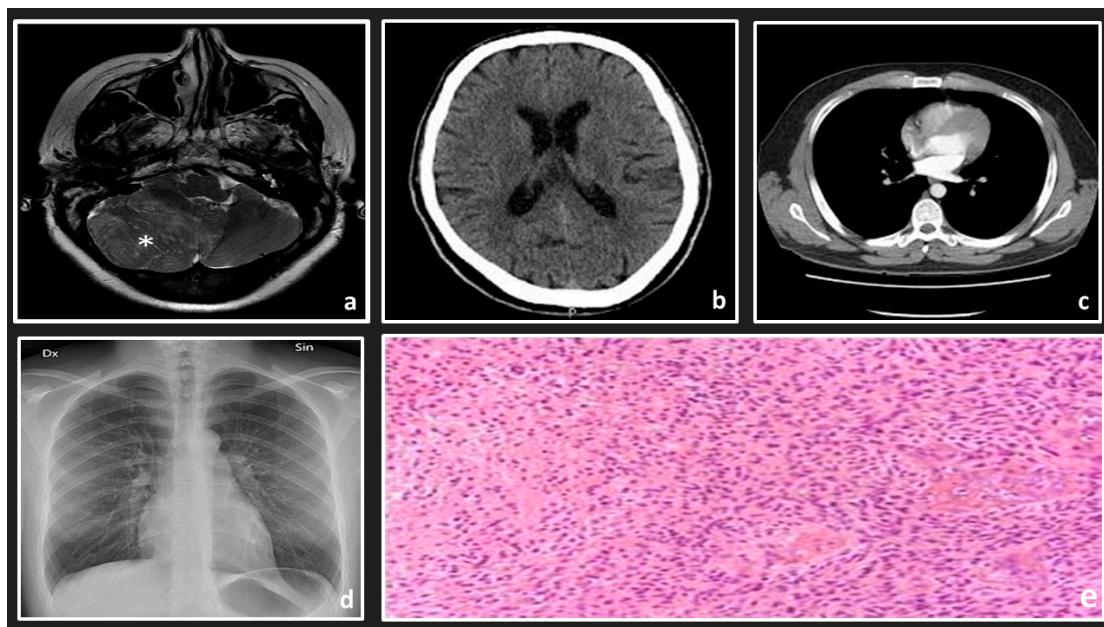
---

## 1. Introduction

Medical Images are a fundamental section of each patient's digital health file. Such images are produced by individual radiologists who are restricted by speed, professional weaknesses, or a lack of practice. It requires decades and reasonable financial resources to train a radiologist. Additionally, some medical care methods outsource radiology confirmations to less economically developed nations, such as India, via teleradiology. A late or incorrect analysis can cause injury to the patient. Thus, it would be beneficial for medical imaging (MI) analyses to be performed by automatic, precise, and effective machine learning (ML) algorithms. MI analysis is a significant research area for ML, in part because the information is somewhat organized and labeled; i.e., this is probable if the patient was examined in a region with good ML systems [1]. That is significant for two reasons. First, with regards to real patient metrics, MI analysis is a litmus check regarding whether ML techniques would, in actuality, improve individual outcomes and survival. Second, it provides a testbed for

human–ML interactions—i.e., how responsive is an individual likely to be to the health changing possibilities being put forward or aided by a nonhuman actor [2]. In recent years, ML has shown significant advances. For a wide variety of applications, including image recognition, medical diagnosis, defect identification and construction health assessments, the potential of this field has also expanded. These new developments in ML are due to many factors, like the creation of self-learning mathematical models that enable computer techniques to execute particular (human-like) tasks based solely on learned patterns, in addition to the increase in the computer power that supports these models' analytical capabilities [3].

There are many imaging types, and their use is becoming more widespread. Types of MI include ultrasound, X-ray, magnetic resonance imaging (MRI), retinal scans, histopathology images (HI), computed tomography (CT), positron emission tomography (PET), and dermoscopy images. Some examples of MIs are shown in Figure 1. Many of these types analyze numerous organs, such as CT and MRI, whereas others are organ-specific, such as retinal and dermoscopy images [4]. The quantity of produced information from each analysis stage differs depending on nature of the MI and the tested organs. HIs are useful for biological studies and to make medical decisions. In addition, they are generally utilized to provide “ground truths” (GTs) for other modalities of MI, such as MRI. A histology slide is a digital record a few megabytes in size, while a magnetic resonance image can be several hundred megabytes. This has a technical effect on how the data is preprocessed and on the architecture design of the algorithm in terms of processor and storage limitations [5].



**Figure 1.** Examples of some medical image types: (a) MRI scan of the left side of a brain; (b) an axial CT brain scan; (c) an axial CT lung scan; (d) chest x-ray; (e) a histology slide with high-grade glioma.

Pathology analyses are traditionally executed by an individual pathologist observing a dyed specimen on a glass slide with a microscope. Lately, efforts have been made to record the whole slide with a reader and save it as an electronic picture, called a whole slide image (WSI) [6].

Digitizing pathology is just one recent development that produces high levels of visible information designed for automated diagnoses. It enables us to see and understand pathologic cell and muscle samples in good quality images with assistance from personal computer tools. It also brings about the possibility of applying image analysis techniques. Such techniques would assist pathologists and support their explanations, such as hosting and grading. Various classification and segmentation methods for HI have already been discussed in this review. We present and compare conventional

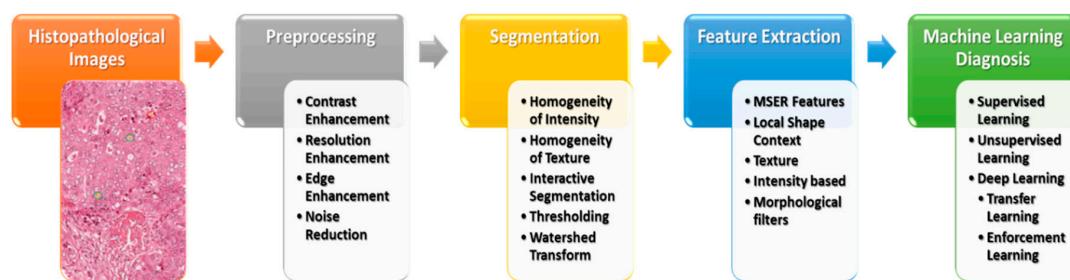
techniques and deep learning (DL) methods to choose the most appropriate method for histopathology issues [7].

Natural microscopic architecture data and their features at nuclei, tissue, and different organ levels could be key to illness expansion and infection treatment analysis. Additionally, to examine and diagnose the histological image of biologic microscopic, pathologists have identified the morphological features of tissue that show the current presence of infection, such as cancer [8].

Some characteristics of disease, such as tumor-infiltrating lymphocytes, might be deduced from HI alone. Additionally, HI analysis, which is called the “gold standard” in many disease diagnoses, is nearly included in all kinds of cancer detection and treatment procedures. HI needs specific analysis with respect to organs and a specific task for the visualization of various tissue components under a microscope. With one or more stains, the sections are dyed. These are staining attempts to uncover cellular elements. The contrast is shown by using counterstains [9].

Efficient ML algorithms are presented and used in HI analysis to help pathologists to acquire a quick, stable, and quantified examination result for a more accurate diagnosis. Many different traditional and deep learning methods support the pathologists in accessing more tissues to determine the internal relationship between the visual images and the specific illness. Additionally, since the ML techniques are generally semi- or fully automated, they are effective, encouraging technical feasibility for histopathology examination within the recent big data age [10].

On the other hand, most of the HI analysis stages are based on mathematical basics. Mathematical operations and functions are applied to all analysis stages, starting from the preprocessing to diagnosis stages to provide an intensive analysis for HIs. Figure 2 illustrates the main phases of a common histopathological images pipeline based on conventional ML techniques. First, HIs are supplied to the system as a 2D array for grayscale images or a 3D array for colored images. Then, the preprocessing stage applies some linear algebra operations on the image array to enhance the image quality. This stage helps to distinguish significant structures from others in the processed images. Third, the segmentation stage is applied to differentiate the cells from other background objects by applying some state-of-the-art mathematical algorithms, such as thresholding, level set, watershed transform, and intensity and texture homogeneity transforms. Fourth, the feature extraction stage extracts the most significant features in the segmented images instead of processing each pixel, which reduces the system’s computation complexity. Besides, most handcrafted features are based on applying some mathematical techniques to detect the changes in the intensity, color, or texture of the pixels. Common derivative techniques are utilized to detect these changes by applying first or second derivatives to pixel values. Finally, the diagnosis stage is applied to classify or cluster the processed images, depending on the extracted features. The classification and clustering techniques are based on applying some mathematical operations that distinguish the processed images based on the extracted features.



**Figure 2.** An overview of the HI analysis pipeline.

Numerous segmentation and classification techniques for primitives tissue in HIs were presented in this respect. For selecting the appropriate HIs analysis method, the various ML methods on HIs are reviewed in this study. In this work, the digital HI analysis, applying different ML algorithms and their issues, are described. The paper is presented to describe the necessity of the analysis procedure for segmentation and classification in computer-aided diagnosis (CAD) systems using HI.

The rest of the article is organized into six sections. Section 2 gives a brief overview of the fundamental histopathological analysis. In Section 3, the conventional approaches for HI analysis are described. Section 4 introduces the use of deep learning techniques in HI analysis. The datasets, the discussion, and tasks of HI analysis are elucidated in Sections 5 and 6, respectively. Limitations and future trends in the HI analysis are introduced in Section 7.

## 2. Histopathological Image Overview

HI has natural and abnormal biological structures, as well as morphological and architectural features defined by pathologists, based on their knowledge. Even given the tissue area, some structures are small, and related patterns typically have high visual appearance variability. In biological systems and anatomy, most visual variability is inherent [11].

Next to obtaining electronic HI via the biopsy test, the guide analysis of images contributes to variability in diagnosis and treatment. To get over this issue, CAD techniques are applied to provide an objective examination of disease. The fundamental steps necessary for applying the CAD examination system appear in Figure 2. This includes electronic image handling methods, such as segmentation, feature extraction, and classification [12].

HI analysis contains the computations executed at various zoom scales ( $\times 2$ ,  $\times 4.5$ ,  $\times 10$ ,  $\times 20$ , and  $\times 40$ ) for multivariate mathematical examination, analysis, and classification. It could be achieved at a lower zoom for tissue stage examination. Demir et al. [13] presented tissue stage and cell stage examination techniques for cancer diagnosis. They examined HI by applying preprocessing, feature extraction, and classification strategies. The new improvement in electronic pathology requirements for the growth of quantitative and automatic digital image examination methods aids pathologists in understanding the number of digitized HIs [14].

### 2.1. Types of ML Systems in HI Analysis

#### 2.1.1. Computer-Aided Diagnosis (CAD)

Many of the searched tasks in electronic HI analysis are CAD systems, which are the pathologists' fundamental functions. The diagnostic method includes the function to map WSI to one of many infection types, indicating a supervised learning function. Considering that the mistakes created by the ML process vary from those created by an individual pathologist [15], the enhancement of classification reliability would be increased by applying the CAD method. CAD could also reduce the instability in understanding and reduce overlooking by analyzing each pixel in WSI. Different related diagnosis functions contain the recognition or segmentation of the region of interest (ROI), such as tumor area in WSI [16,17], rating of immunostaining [18,19], cancer phase [20,21], mitosis recognition [22,23], gland segmentation [24–26], and quantification of general intrusion [27].

#### 2.1.2. Content-based on Image Retrieval (CBIR)

CBIR retrieves pictures related to a query picture. In electronic pathology, CBIR methods help in several scenarios, especially in examining, training, and studying [28–30]. For example, CBIR methods could be utilized for academic applications and novice pathologists to recover appropriate instances of HI of tissues. Additionally, such methods would also be useful to skilled pathologists, especially while detecting uncommon cases. Because CBIR certainly does not need tag data, unsupervised learning could be utilized [31]. Not just precision, but additionally high-speed research of related pictures from several pictures are needed in CBIR. Thus, numerous approaches can reduce picture feature dimensionality—such as primary element examination and small bilinear combination [32]—and quickly estimated the closest neighbor searches [33].

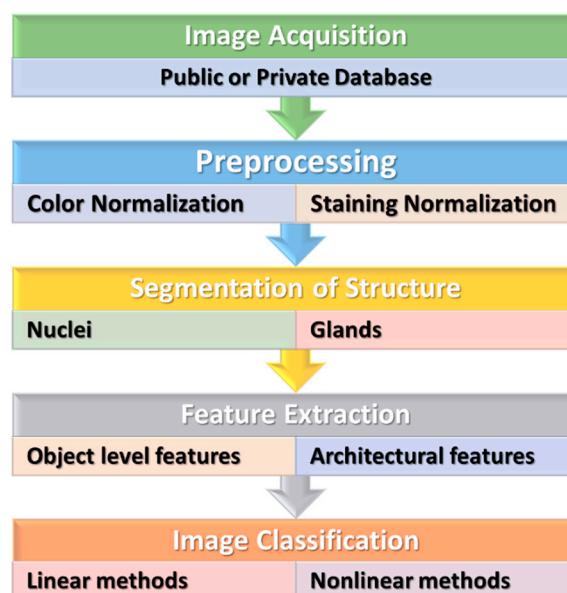
### 2.1.3. Finding New Clinicopathological Associations

Traditionally, several essential discoveries regarding diseases, such as tumor and contagious conditions, have already been produced by pathologists and analysts. They cautiously and carefully examined pathological specimens. For example, pathologists analyzed the gastric mucosa of individuals with gastritis in [34]. Efforts were made to link the morphological options that come with cancers using their medical behavior. For instance, tumor grading is essential in a patient's diagnosis and in preparing treatment for many kinds of cancer, such as breast and prostate cancer.

There is a noticeable development in the digitization of clinical data, which later improved the genome evaluation technique. Therefore, a wide range of electronic data, such as genome data, electronic pathological images, MRI, and CT scans, are now accessible [35]. By examining the connection between these imaging modalities, new hospital pathological associations, such as the connection involving morphological quality and somatic cancer mutation, are available [36]. CAD techniques could be subdivided into conventional ML and DL methods, illustrated in more detail in the next few sections.

## 3. Conventional Machine Learning Methods

CAD systems played an essential role and have become an important research topic in HI and diagnostics. Various image processing techniques were applied to examine the disease's diagnosis and prognosis for these HIs. Various image processing and computer vision (CV) techniques have been implemented for gland and nuclei segmentation, cell kind recognition, or classification to extract quantitative measurements of disease characteristics from HIs and automatically assess whether or not a disease exists inside examined samples. It could help to determine the degree of seriousness of the disease, whether present in the sample. Conventional ML methods often contain a few steps to manage HI, as shown in Figure 3. Each step is illustrated in the following sections.



**Figure 3.** The conventional machine learning methods for HI.

### 3.1. Preprocessing

Preprocessing could recompense for variations between images, which can vary in color, staining, and other problems, such as noise, which are usually due to the scanning procedure. The gross sections are made with wax to analyze the tissue's architecture and components under the microscope and colored with one or more stains. Pathologists use staining to isolate cellular components for the diagnosis of structural as well as architectural tissue analysis. Hematoxylin–Eosin (H&E) staining is

most commonly utilized, and it separates the connective tissue, cytoplasm, and nuclei. Nuclei are stained blue by Hematoxylin, while connective tissue and cytoplasm are stained pink by Eosin. DAB, immune-histochemistry stains, etc. are the other stains. The consistency of the features extracted from the image directly affects classification performance. Thus, it is essential to define the proper conditions under which the image preprocessing techniques will work as the first job. Noise and various illumination fluctuations are detrimental to image processing techniques. If those negative factors are eliminated, it will improve performance. Pre-processing imaging techniques are well adapted to this mission. Pre-processing methods control changes in image brightness and contrast and eliminate noise.

### 3.1.1. Staining Normalization

HI could have powerful color variations due to various scanners, various staining techniques, and sample age. An efficient color calibration between samples is difficult to accomplish [37]. Hence, color normalization is needed in most of the processing scenarios. Deconvolution-based methods and histogram-based methods are examples of color normalization [38]. Anghel et al. [39] suggested improving stain normalization in low-quality WSIs to increase ML pipeline accuracy. They used an ML pipeline based on convolutional neural networks (CNNs), which classifies pictures to detect prostate cancer, to demonstrate the robustness of this new normalization process. This system makes it possible to pre-process massive datasets and is a crucial requirement for any biomedical imagery learning computer.

### 3.1.2. Color Normalization

Color normalization is required for bright and fluorescent HI analysis. This process decreases the variations in samples of tissue because of variance in conditions of scanning and staining. There are different techniques for the color normalization of HI, such as the Reinhard approach, descriptor of stain color, and histogram specification [40]. For HI research, MIAQuant [41] was stained by different approaches and obtained with various instruments. The machine automatically extracts and quantifies markers with various colors and types and, for the visual comparison of their positions, aligns the contiguous tissue slices, stained by multiple markers. MIAQuant segments markers efficiently and quantifies them by integrating clear and effective imaging techniques with precise colors from histological images. MIAQuant aligns and measures a picture with contiguous (serialized) parts of the cloth, where the markers are covered with different colors so that the markers can be visually comparable. Its successful findings in biomedicine have inspired us to increase the capacity to communicate different marker positions and, finally, neighboring serialized pieces of tissue. Easy, efficient, and effective processing, pattern detection, and supervised teaching techniques with their improved framework, called MIAQuant-Learn [42], enable you to personalize marker segmentation by all colors.

The quality of HI is the parameter to determine that it can be the most remarkable approach for color normalization. Metrics of quality, such as the structural similarity index metric (SSIM), contains three factors (contrast, luminance, and structural) [43]. These parameters are given in Equations (1)–(3), respectively.

$$N(x, y) = \frac{2 \sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (1)$$

$$M(x, y) = \frac{2 \bar{X} \bar{Y} + c_1}{\bar{X}^2 + \bar{Y}^2 + c_1^2} \quad (2)$$

$$R(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (3)$$

where  $\bar{X}$  and  $\bar{Y}$  are means of origin and processed image, respectively.  $\sigma_x$  and  $\sigma_y$  are standard deviation,  $\sigma_{xy}$  is the correlation coefficient between the processed and the source image.  $c_1$ ,  $c_2$ , and  $c_3$  are constants that could stabilize SSIM if nearing a zero value. By using Equations (1)–(3), the SSIM is derived from Equation (4). The SSIM value is 0 to 1. The better approach is color normalization, when the value is near to 1.

$$\text{SSIM} = \left( \frac{2 \bar{X} \bar{Y} + c_1}{\bar{X}^2 + \bar{Y}^2 + c_1^2} \right) \left( \frac{2 \sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right) \quad (4)$$

### 3.2. Recognition and Segmentation of Structures

One of the main tasks in HI analysis is image segmentation, and it has been applied to solve a wide variety of issues. Image segmentation, in its entirety, similar to clustering, is an unplaced issue as defining a meaningful segment can vary from task to task or even from image to image. For this purpose, the application domain must be aware of the segmentation algorithms, either by taking custom features or algorithmic methods into account or by learning from vast volumes of knowledge [44].

The existence of pathology and the number and the morphological features of detailed textures, such as nuclei and glans, are essential variables to analyze the existence and intensity of pathology—for example, colorectal [45], prostate [46], and breast [47] cancer.

#### 3.2.1. Nuclei and Cells

Nuclei would be the main organelles of a eukaryotic cell, comprising the majority of cell DNA. Nuclei examination often requires recognition, segmentation, and separating overlaps. Recognition of seed factors in nuclei is needed by several segmentation and checking techniques [48]. Several methods had been proposed in the review for nuclei recognition, involving techniques predicted on Euclidean range chart peaks [49], Hough change (recognizing seed factors for circular formed textures, requesting extensive computation) [50], Laplacian of Gaussian filters [51], and radial symmetry [52]. Several methods were shown to represent accurate segmentation. The techniques based on thresholding and morphological procedures are appropriate on a standard background [53]. They may not, however, be powerful in measurement, form, and structure change. Effective shape forms could mix picture attributes with nuclei form types [54]. However, they depend on seed factors. Different techniques were predicated on gradients in polar [55] and graph reductions [56]. Ta et al. [53] suggested an approach, dependent on regularization graphs. The method's specificity was to utilize graphs as image confidential modeling at various grades (areas or pixels) and various component relations, such as grid graph. Dependent on Voronoi's diagrams, they suggested a graph reduction technique for nucleus segmentation of HIs for serious cytologic and breast cancer. A pseudo metric  $\delta: V \times V \rightarrow R$  is illustrated as

$$\delta(u, v) = \min_{p \in P_G(u, v)} \sum_{i=1}^{m-1} \sqrt{w(u_i, u_{i+1})} (f(u_{i+1}) - f(u_i)) \quad (5)$$

where a weight function between two pixels is  $w(u_i, u_{i+1})$ , and a set of paths connecting two vertices is  $P_G(u, v)$ . Taking into account the set of  $K$  seeds  $S = (s_i \subseteq V)$ , where  $i = 1, 2, \dots, K$ , the energy  $\delta: V \rightarrow R$  presented the metric  $\delta$  for all the seeds of  $S$ , which can be presented as:

$$\delta_s(u) = \min_{s_i \in S} \delta(s_i, u), \quad \forall u \in V \quad (6)$$

The zone  $z$  of control (known as the Voronoi cell) of the seed  $s_i \in S$  is the set of vertices nearer to  $s_i$  than any other seeds related to the metric  $\delta$ . It can be defined,  $\forall j = 1, 2, \dots, K$  and  $j \neq i$ , as

$$z(s_i) = u \in V : \delta(s_i, u) \leq \delta(s_j, u) \quad (7)$$

Then, the energy distribution of the graph is the set of powerful zones  $Z(S, \delta) = \{Z(s_i), \forall s_i \in S\}$ , for a given set of seeds  $S$  and a metric  $\delta$ .

### 3.2.2. Glands

It is organs that are shaped by an ingrowth from the epithelial surface. Techniques in thresholding and area growing could recognize nuclei and lumen, which can be applied to initial seed factors for area growing [57]. Segmentation predicated on polar ordinates (the middle of the gland) was performed on the benign and malign gland [48].

Rittscher et al. [58] caught some bright pixels participating in a standard distribution form. Their technique applied three features. The first was the intensity of fluorescent emission. The others were derived from curve descriptors, which could be calculated from eigenvalues of the Hessian matrix. The eigenvalues ( $\lambda_1(x, y) \leq \lambda_2(x, y)$ ) of the image  $I(x, y)$  encode the curve data of the image. They give helpful cues for shape detection, such as structures of the membrane. However, the eigenvalues are influenced by the brightness of the image [6]. Equations (8) and (9) represent two features, which are autonomous from the image's brightness. These features are known as normalized-curve index and shape index, respectively.

$$\phi(x, y) = \tan^{-1} \frac{\sqrt{(\lambda_1(x, y))^2 + (\lambda_2(x, y))^2}}{I(x, y)} \quad (8)$$

$$\theta(x, y) = \tan 2(\lambda_1(x, y), \lambda_2(x, y)) \quad (9)$$

This segmentation, based on normalized-curve index and shape index divides an image's pixels into three sets: foreground, indefinite, and background. The indefinite set covers all of the pixels which are not involved in the other two sets. From these sets, the intensity distribution of foreground and background and log-likelihood intensity are derived.

### 3.3. Feature Extraction

HI was examined by applying several descriptors based on the information of domain experts. Analysis requirements are represented primarily with cytological phrases (i.e., glands, nuclei) and its involvement in the malignant and benign surface. Consequently, many papers dealt with the object stage (applying segmented item attributes) and object connection stage (applying structural attributes). Tumors, such as ductal carcinoma and lobular carcinoma, have an abnormal growth of epithelial cells in these structures. The abnormal growth of tissue, representing a tumor, may result in a large number of nucleus cells or a high number of mitotic cells in a small area. HI captures this function, but it captures other healthy tissues in addition to the nucleus, which can be seen in images of benign tumors. Stroma is a kind of tissue that can be seen in malignant and benign images with the same characteristics. The classification method could be enhanced by choosing more appropriate patches. Histopathological considerations remain paramount in this regard. There are well-known considerations, such as the size of the tumor, the histological shape and subtype, the nature of the sign, circular morphology and degree of differentiation, and the presence of vascular lymph invasion and the involvement of lymph nodes. We have gone from a greater understanding of these causes in recent years, identifying significant factors, such as tumor budding and lymphocytic infiltration. The prognostic importance of resection margins has also been assessed over the last two decades—particularly circumferential margins. Some patients are also notable with histological features associated with various molecular and genetic markers, including KRAS, BRAF, and microsatellite instability. The feature can be divided into literature level artifacts and structural features. Object-level characteristics are characteristics correlated with the nucleus size and shape. Features of the structure describe topological characteristics based on graph theory.

### 3.3.1. Object-Level Characteristics

Object-level characteristics rely firmly on regarded items (often gland or nuclei) and segmentation methods [6]. These characteristics are appropriate at any resolution. However, generally, they are produced by high-resolution pictures. Object-level characteristics were generally produced to each shade channel and could be collected into shape characteristics, such as region [59]. In the pre-segmentation procedure based on the unregulated medium-shift cluster, Kuse et al. [60] applied a feature extraction method. Thresholds limit the color variation to the image section. After this stage, the kernel is formed, and the contour and area constraints are removed from the overlap. Gray-level co-occurrence Matrix (GLCM) texture functions are finally extracted from a sectioned picture utilizing a classifier specified by support vector machines (SVM). Caicedo et al. [24] merged seven approaches of extraction of features and construct a kernel-based representation of data for each form of function. Inside the SVM classifier, kernels are used to find similarities between data and enforce a content retrieval mechanism.

### 3.3.2. Structural Characteristics

Structural characteristics are primarily based on graphs. They are comprised of nodes that can be linked by arcs. Lately, characteristics based on the graph were investigated frequently because they are suitable to characterize tumor structure. Three types of brain cancer classification (inflame, health, maligning) were conducted by Demir et al. [13]. They utilized a total weighted graph. Chekkoury et al. [61] made a hybrid between features based on texton and morphologic system to breast cancer. Doyle et al. [62] applied a mix of characteristics for prostate cancer. They show 90% precision in characterizing cancer and benign tissue. To handle the multi-resolution properties of HIs and emulate pathologists' approach to analytics, multi-resolution methods have been proposed. In many resolutions, the Gaussian pyramid method portrayed the pictures. Features for every level have been extracted separately and labeled as image tiles. Color and texture characteristics are typically utilized at low resolutions. The medium-scale architectural arrangement of glands and nuclei at high resolutions may be discriminatory.

## 3.4. Classification

Techniques for classifying their goal determine the group of recent observations between some classes based on marked training group. Regarding the function, anatomical composition, characteristics, and the preparation of tissue, the classification differs. DiFranco et al. [63] classified the prostate tissue into seven classes: benign hyperplasia, inflammation, Gleason rank 3 and 4, and intraepithelial neoplasia. They acquired 83% accuracy based on sum average, contrast, connected histogram characteristics, and entropy.

Huang et al. [64] proposed a method to enhance hepatocellular carcinoma classifying, which used a subset of feature selection with a support vector machine based on a decision graph for every decision node. Alexandratou et al. [65] presented a literature review for prostate cancer, illustrating good cancer recognition. The overview of conventional HI analysis methods is summarized in Table 1.

**Table 1.** Overview of HI analysis for different conventional methods.

| Study                     | Organ      | Method  | Results   | Problem with Method   |
|---------------------------|------------|---|---|---|
| Basavanhallya et al. [66] | Breast     | Hierarchical Normalized cut   | 89% accuracy of segmentation  | Discovers false-positive errors because of lumen presence.  |
| Khadi [67]                | Meningioma | Classification of texture applying fractal features   | 92.5% accuracy applying individual texture measurement to meningioma tissue                               | Misclassification results because of the non-uniformity cell construct.   |
| Demir et al. [68]         | Colon      | Object graph approach for segmentation  | 87.59% accuracy of compatible images  | Requires variable optimization that reduced results for segmentation  |
| Tosum et al. [69]         | Breast     | Diagram run-length models to segmentation of the image  | The novel descriptor of texture for unsupervised classification had 99.0% accuracy for gland segmentation | Complexity relies on the number of primitives in picture  |
| Chekkoury et al. [61]     | Breast     | A novel hybrid between features based on texton and morphologic system  | 86% classification accuracy combining textural features and morphometric                                  | The effects of image compression on classification accuracy   |
| Doyle et al. [62]         | Prostate   | mix of characteristics  | 90% accuracy in characterizing cancer and benign tissue   | Limited Feature Set   |
| DiFranco et al. [63]      | Prostate   | classified into seven classes (benign hyperplasia, inflammation, Gleason rank 3,4, and intraepithelial neoplasia) | 83% median accuracy based on some characteristics like sum average, contrast                              | The computation time required for reading and writing data to and from disk, particularly in feature extraction |

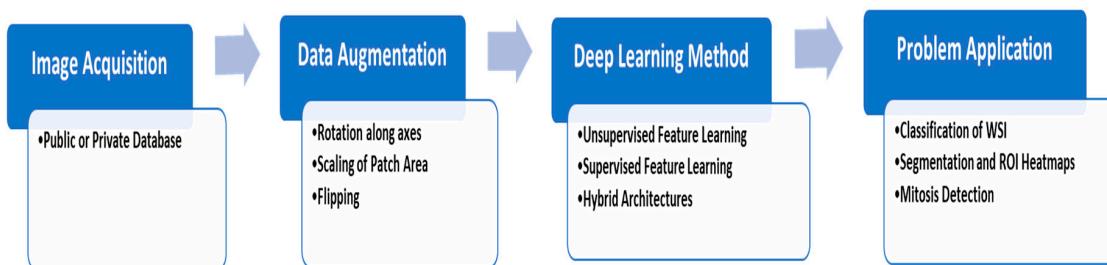
#### 4. Deep Learning Methods

Recently, DL techniques have often been studied in the effective form of ML methods. Within the last few years, DL techniques outperformed traditional ML methods in varied fields, such as CV, natural language processing (NLP), biomedical fields, and automated analysis for HI [7]. DL methods in the CV are derived from the structure levels for nonlinear transformations on natural input pixels. This structure formed significantly abstract representations, which could be realized in a hierarchical style [70]. A typical instance of a commonly applied structure is the CNN [71].

Multiple criteria can be considered when using the DL techniques to deal with histopathology, since accomplishing the method is partly due to the task-species setting. Among the principal features of HI is that appropriate styles be determined by the magnification stage. The key factors are the size of the patch given to the network, the localization of parts in the image where appropriate histopathology originals can be found, and the homogeneity of staining for WSI [72]. The network structure represents an important position, while many studies keep predefined system structures, as illustrated in Figure 4.

The majority of the DL techniques for localizing, classifying, and segmenting HI are somewhat recent. Deep neural techniques are stated in the new literature of HI analysis, such as [6,13,36]. For example, Irshad et al. [48] were the first mentioned in a review. The critical patterns from an exhaustive analysis of different nuclei identification, segmentation, and classification approaches utilized in HI, specifically in H&E staining protocols, were described and discussed in this review. Ciresan et al. [56] presented one of the first significant efforts to utilize the deep method in mitosis

recognition for HI analysis. Arevalo et al. [73] presented a hybrid illustration method to the basal cell of carcinoma areas and utilized a topographic unsupervised technique and a case of characteristic illustrations. They increased the classifier's efficiency by 6% regarding traditional structure-based discrete cosine transform (DCT). Nayak et al. [74] presented an alternative method for the unsupervised Boltzmann technique for understanding image signatures. They classified images of the cancer genome atlas (TCGA) for apparent cell-kidney cancer and glioblastoma variform. The last stage was created utilizing the classifier of multi-class support vector machines (SVM) techniques.



**Figure 4.** The typical deep learning steps for HI analysis.

Malon et al. [75] proposed a novel mix of features. The authors compiled a feature set of the fundamental nucleus and cytoplasm pixel statistical measurements and combined them with a CNN classifier. Their method increased the efficiency when comparing to the characteristics of handcrafted methods. Xu et al. [76] created an approach that handled marked cases. Numerous examples of the learning platform were presented in this technique, where the colon HI classifier was developed. The researchers presented a thoroughly supervised method and also weakly-marked one. Hou et al. [77] presented a similar strategy by applying numerous cases to understand how to categorize low-grade glioma and glioblastoma images in TCGA. The technique used three phases. First, it understood the masks of discriminative parts utilizing CNN form with few picked discriminative areas. It then created a patch level forecast, applying CNN. Finally, the counting of the class was generated.

Arevalo et al. [78] proposed a stacked form that revealed the most significant features when mixing characteristics from two-layered topographic independent component analysis (TICA) around patches for finding basal cell cancer. They presented an electronic discoloration technique because characteristic detectors are weighed with classification likelihood to spotlight parts, which can be most linked to carcinoma. This technique accomplished 99% of the area under the curve (AUC) for 100,000 patches. Hang et al. [79] were dependent on the understanding book of 1024 characteristics to classify apparent cell cancer and also glioblastoma multiforme. The book was constructed with a stacked unsupervised technique utilized in the spatial chart corresponding platform with the SVM classifier's final step. Some new studies were interested in the classifying of the H&E patch. Han et al. [80] provided a novel deep unsupervised technique for glioblastoma multiforme characterization. They can distinguish two critical phenotypic subtypes at various survival shapes, utilizing the produced unsupervised characteristics. Noel et al. [81] applied a group of 30,000 patches produced from the WSI of the International Conference on Pattern Recognition (ICPR) contest to recognize breast carcinoma, categorizing every pixel utilizing CNN mitosis, stoma, and lymphocytes. They reached 90% accuracy, indicating a better WSI primitive classifier, which could enhance classification performance.

Romo-Bucheli et al. [82] served a CNN form with prospect patches for calculating tubule associateship likelihood to measure tubule nuclei, which was related to high-low chance classes decided by the Oncotype DX test. DL stains' application distinctive from H&E continues to be not adequately researched, such as immunohistochemistry. Chen et al. [83] presented the recognition of immune cells using seven-layer CNN with areas of nature colors from RGB channels, displaying markers of immune cells. They compared the efficiency of their recognition method with pathologists' efficiency, attaining a 99% correlation coefficient. A review of deep-neural models, developed for HI analysis, was presented by Srinidhi et al. [84].

Sumi et al. [85] proposed a deep spatial fusion network that manages the dynamic construction of discriminative characteristics over patches. In the high-resolution histology picture, it also learns to change the bias on patch-wise prediction. To extract characteristics from the cellular level to the tissue level, Patch wise InceptionResNetV2 is used. This approach is used to analyze the spatial interaction between patches. Compared to previous CNN experiments using different architectures, better performance is given by their proposed system. This work needs to be expanded to include other networks that effectively examine more malignant tumor types other than glioblastoma and oligodendrogloma. Maximal tumor resection is especially necessary for larynx surgery, thus maintaining adjacent healthy tissue. Therefore, accurate and swift intraoperative laryngeal histology is vital for optimal surgery results. Zhang et al. [86] hypothesized a DL stimulation of Raman scattering microscopy (SRS) that could automatically and precisely diagnose new, unprocessed surgical specimens with laryngeal squamous cell cancer without fastening, separation or staining. First, they compared 80 pairs of adjacent SRS and regular frozen parts to determine their concordance. They then applied SRS imaging to 45 patients' fresh surgical tissues based upon a DL model for automatically producing histologic results. They also applied the main diagnostic features.

Pathology's scientific function is to diagnose diseases to classify differences at cell structures level, such as nucleolus and cytoplasm, tissue (i.e., cell community with complicated structures), and organs that give rise to patient symptoms. It was found that damaged or unresolved cells do not die, and uncontrolled growth is seen by clinical pathology framework and histopathology methods, explaining cancer cells' mass production. Cancer cells also migrate through the blood and lymph systems and cross borders to another body area, reproducing the uncontrolled growth cycle method. This cancer cell phase is called metastatic spreading or metastasis that leaves one area and grows in another part of the body. Breast cancer could be detected by a histopathological methodology. This diagnosis can be made using different ML Models, and DL-based CNN Models. Agarwal et al. [87] showed that CNN models provided significant accurate results in the comparison of ML models, such as U-Net [88], improved precision, and high-performance segmentation. In conjunction with very low-cost consumer graphics processing units, large images can therefore be processed rapidly.

Approaches that rely on Generative Adversarial Networks (GANs) are likely to minimize the need for large volumes of manual notations. Not only have recent innovations enhanced initiatives but so have new technologies. Now, unattended techniques may carry out various tasks for which supervised methods are indispensable. The latest state-of-the-art advances in histopathological images of GANs were summarized in [89]. The overview of the discussed studies is summarized in Table 2.

**Table 2.** Overview of supervised and unsupervised learning models based on DL techniques.

| Study                      | Organ            | Staining | Potential Usage                         | Method  |
|----------------------------|------------------|----------|---|---|
| <b>Supervised Learning</b> |                  |          |   |   |
| Litjens et al. [90]        | different tissue | H&E      | Prostate and breast carcinoma detection | Convolutional Neural Network based on pixel classifier  |
| Nagpal et al. [91]         | Prostate         | H&E      | Anticipating Gleason indicator          | CNN based on sectional Gleason model classifier + k-nearest neighbors (KNN) based on Gleason grade anticipation                 |
| Zhao et al. [92]           | Breast           | H&E      | Metastasis Detection + classification   | Characteristic pyramid collecting based on the fully convolutional network (FCN) system with the synergistic training technique |

**Table 2.** *Cont.*

| Study                        | Organ            | Staining                              | Potential Usage                                | Method  |
|------------------------------|------------------|---------------------------------------|--|---|
| Xing et al. [93]             | different tissue | H&E,<br>Immunohistochemistry<br>(IHC) | Segmentation of nuclei                         | CNN + selection based on sparse form Pattern  |
| Gu et al. [94]               | Breast           | H&E                                   | Tumor detection                                | U-Net based on multiple resolution model with multiple encoders and a singular decoder system   |
| Tellez et al. [95]           | Breast           | H&E                                   | Detection of Mitosis                           | Train of Convolutional Network applying H&E registered to PHH3 slides as a reference  |
| Wei et al. [96]              | Lung             | H&E                                   | Histological subtypes of lung gland classifier | ResNet-18 on the basis of patch classification  |
| Song et al. [97]             | Cervix           | Papanicolaou (Pap),<br>H&E            | Cells Segmentation                             | Multiple level CNN system   |
| Agarwalla et al. [98]        | Breast           | H&E                                   | Segmentation of tumor                          | CNN and 2D- Long short-term memory (LSTM) to representing training and context collecting   |
| Ding et al. [99]             | Colon            | H&E                                   | Glands segmentation                            | Multiple level FCN network with a high-resolution section to avoid the lost in highest pooling layers   |
| Bejnordi et al. [100]        | Breast           | H&E                                   | Invasive Carcinoma detection                   | Multiple level CNN which first determines tumor-associated stromal modifications and more categorize into normal/benign versus invasive carcinoma |
| Seth et al. [101]            | Breast           | H&E                                   | Ductal carcinoma in-situ (DCIS) segmentation   | Compared UNets learned in many resolutions  |
| <b>Unsupervised Learning</b> |                  |                                       |  |   |
| Xu et al. [102]              | Breast           | H&E                                   | Segmentation of nuclei                         | Stacked sparse autoencoders   |
| Bulten and Litjens [103]     | Prostate         | H&E                                   | Tumor classification                           | Convolutional adversarial Autoencoders  |
| Hou et al. [104]             | Breast           | H&E                                   | Segmentation and detection of nuclei           | Sparse autoencoder  |
| Sari and Gunduz-Demir [105]  | Colon            | H&E                                   | Feature extraction and classification          | Restricted Boltzmann + clustering   |
| Gadermayr et al. [106]       | Kidney           | Stain agnostic                        | Object of interest segmentation in WSIs        | CycleGAN + UNet segmentation  |
| Gadermayr et al. [107]       | Kidney           | Periodic acid-Schiff (PAS), H&E       | Glomeruli segmentation                         | CycleGAN  |

## 5. Datasets

The size of the datasets given to researchers for training and testing their methods has dramatically increased in the latest challenges. There is a set of public databases in the electronic pathology subject that include manual annotations for HI, as listed in Tables 3 and 4 [108]. They might help the examination objectively. Slide issue (stain) and image issue (image resolution, zoom level) are similar. However, all these databases are targeted to specific diseases. These databases do not handle several tasks. Additionally, there are many high scale HI datasets, which include WSIs of high resolutions.

TCGA [33] includes around 10,000 images from different types of cancer. Genotype-Tissue Expression (GTE) [109] includes around 20,000 WSIs from different tissues. The Stanford Tissue Microarray Database (TMAD) is available for researchers to access images of microarrays for tissue. It provides images of archiving 349 distinguished probes on 1488 microarray slides of tissue [110]. The CAMELYON dataset is a collection of WSI tissues for the sentinel lymph node. It contains CAMELYON16 and CAMELYON17 challenges that include 399 WSI and 1000 WSI, respectively. The data are currently accessed via registration on the CAMELYON17 website [111]. The Breast Cancer Histopathological Image (BreakHis) contains 9109 macroscopic images for the tissue of the breast tumor obtained from 82 patients in various magnifying factors (40X, 100X, 200X). Up to now, it includes samples of 2480 benign and 5429 malignant WSIs [112].

**Table 3.** Some common downloadable WSI databases.

| Datasets               | No Slides     | Staining | Diseases                              |
|------------------------|---------------|----------|---------------------------------------|
| TCGA [33,113]          | 18,462        | H&E      | Cancer                                |
| GTE [109]              | 25,380        | H&E      | Normal                                |
| TMAD [110,114]         | 3726          | H&E/IHC  | various tissue                        |
| TUPAC16 [115]          | 821 from TCGA | H&E      | Breast cancer                         |
| Camelyon17 [111]       | 1000          | H&E      | Breast cancer (lymph node metastasis) |
| Köbel et al. [116,117] | 80            | H&E      | Ovarian carcinoma                     |
| KIMIA Path24 [118]     | 24            | H&E/IHC  | various tissue                        |

**Table 4.** Some publicly available hand-annotated histopathological images.

| Datasets                        | No of Images | Staining | Organs           | Potential Usage                       |
|---------------------------------|--------------|----------|------------------|---------------------------------------|
| KIMIA960 [119,120]              | 960          | H&E/IHC  | Different tissue | Classification                        |
| Bio-segmentation [121,122]      | 58           | H&E      | Breast           | Classification                        |
| Bioimaging challenge 2015 [123] | 269          | H&E      | Breast           | Classification                        |
| GlaS [124]                      | 165          | H&E      | Colorectal       | Gland segmentation                    |
| BreakHis [112]                  | 7909         | H&E      | Breast           | Classification                        |
| Jakob Nikolas et al. [120,125]  | 100          | IHC      | Colorectal       | Detection of blood vessel             |
| MITOS-ATYPIA-14 [126]           | 4240         | H&E      | Breast           | Detection of mitosis, classification  |
| Kumar et al. [119,127]          | 30           | H&E      | Different cancer | Segmentation of Nuclear               |
| MITOS [20]                      | 100          | H&E      | Breast           | Detection of mitosis                  |
| Janowczyk et al. [128,129]      | 374          | H&E      | Lymphoma         | classification                        |
| Janowczyk et al. [128,129]      | 85           | H&E      | Colorectal       | Segmentation of gland                 |
| Ma et al. [130]                 | 81           | IHC      | Breast           | TIL analysis                          |
| Linder et al. [131,132]         | 1377         | IHC      | Colorectal       | Segmentation of epithelium and stroma |

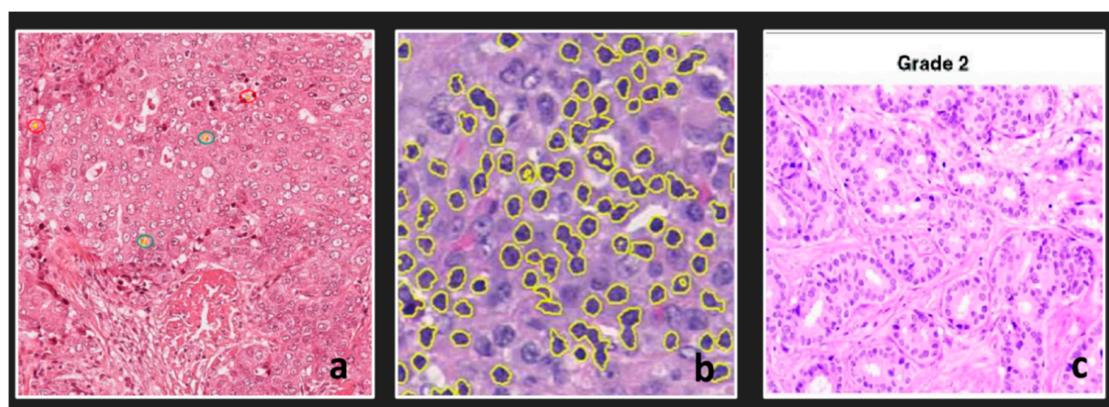
## 6. Discussion and Histopathological Tasks

Since HI analysis is inherently a cross-disciplinary area, this review has stated that ongoing research is anticipated to have an obvious and tangible impact on automated HI analysis techniques. This paper reviews the recent state of the art CAD techniques for HI. This review also briefly describes the

development of histopathology analysis and its problems. Recently, DL outperformed state-of-the-art techniques in various MLs for HI analysis tasks, such as recognition, classification, and segmentation. DL's merit, compared to other forms of learners, is their ability to acquire the performance as well or better than a human's performance. Currently, DL and WSI are revolutionizing the CAD of histopathology, and soon they could help reduce pathologists' workload in most simple tasks. This would allow pathologists to focus on challenging cases and lead to a deeper comprehension of pathologic procedures via ML techniques. More applications of HI analysis using ML techniques have been introduced in this review [108]. Most of the research developed in the field of HI analysis is addressed for some specific tasks.

#### *Tasks for Histopathology Image*

Open objective problems targeting issues in HI analysis were presented recently, such as in other medicinal imaging fields. A benefit of trying various methods on an unchanging dataset and in exact issues is the target comparison of advantages and constraints of literature methods. Especially in a sample of pathology, consuming time, and challenging issues of searching WSIs for appropriate tissue basics, such as nuclei and mitosis, could be enhanced. Choosing the most appropriate method to aid and advance the visible model of slides could help pathologists concentrate on significant issues when analyzing the mentioned studies. The difficulty of issues has improved recently. The objectives of problems could be gathered into three major issues, as shown in Figure 5.



**Figure 5.** Challenges of HI analysis: (a) Mitosis detection, (b) Segmentation, and (c) Grade classification of a tumor.

- **Recognition of Mitosis**

Recognize mitosis contained in large power domains: There is a powerful relationship with the aggression of carcinoma and faster cell separation in extra mitosis. An essential part of HI tasks is the proper selection of evaluation metrics. A task of mitosis recognition utilizes the F1-score as the best metric to evaluate the participant techniques. The F1-score is calculated using Equation (10).

$$\text{F1-score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

where Precision = TP/(TP + FP) and Recall = TP/(TP + FN). The threshold of maximum Euclidean distance from centroid for considering the mitotic event, as TP, was estimated to less than 7.5–8  $\mu\text{m}$ .

- **Segmentation of structure**

The segmentation process localized and outlined the border of particular tissue architectures—for example, nuclei of cell or gland. Various kinds of tissue structures have been

aimed to structure segmentation in HI. Automatic segmentation system output is typically evaluated by measuring some standard objective parameters, such as the mean boundary distance, the Dice coefficient, and Hausdorff Distance (HD). HD is one of the most insightful and useful metrics, since it measures the greatest segmentation error. For the two datasets, X and Y, the one-sided HD from X to Y is defined as:

$$\text{hd}(X; Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|^2 \quad (11)$$

and similarly, for  $\text{hd}(Y; X)$ :

$$\text{hd}(Y; X) = \max_{y \in Y} \min_{x \in X} \|x - y\|^2 \quad (12)$$

The bidirectional HD between these two sets is then:

$$\text{HD}(X; Y) = \max(\text{hd}(X, Y), \text{hd}(Y, X)) \quad (13)$$

#### • Classification of images

First, one must find the characteristic set of features for a specific class of tissue, potentially taking into account the primitives of the underlying tissue. In histopathology, various evaluation techniques for ranking the classifications methods have been used. Various evaluation methods for rating the classification methods include

- For nuclear atypia rating, a points-based scheme is used.
- The region under the curve of the receiver operating characteristic (ROC) to classify slides of lymph node comprising metastasis or not.
- Approval with ground truth calculated with Spearman's correlation or quadratic weighted Cohen's kappa to grade cancer.

Equation (14) is used to calculate quadratic weighted Cohen's kappa.

$$K_w = 1 - \frac{\sum_{i,j} w_{i,j} p_{i,j}}{\sum_{i,j} w_{i,j} e_{i,j}} \quad (14)$$

$$e_{i,j} = p_{i,j} q_{i,j} \quad (15)$$

where  $w_{i,j}$  represent weights,  $p_{i,j}$  observed probabilities, and  $q_{i,j}$  expected probabilities [133]. Though these issues were split into various sets due to their explanation, they could be mixed or regarded as a preliminary phase to another issue. For example, the WSI grading technique of carcinoma could begin by classifying the image as a tumor tissue. Next, the carcinoma was segmented. Finally, grade WSI dependent on the count of mitosis in the ROI of cancer.

## 7. Limitations and Future Trends

Digital HI recognition is an appropriate issue for ML because pictures themselves include data adequate for diagnosis. Issues in the analysis of digital HI applying ML is mentioned in this review. Because of reasonable efforts produced up to now, these issues being overcome, but there is space for enhancement. Many of these issues are probably resolved when a large amount of well annotated WSIs becomes obtainable. Collecting WSIs from different institutions to note them the exact conditions and creating this information public will be adequate to improve the growth of more advanced electronic HI analysis. Lastly, some possible future issues for the study are recommended, which have not been adequately researched.

#### • Novel Objects Discovery

For instance, unexpected items, irregular organization, uncommon tumor (not contained in the training stage), and aliens' bodies might exist in real diagnostic conditions. However, one can use a

discrimination framework containing CNN classes, such as items among the predefined classes [134]. To solve the issue, the recognitions of outlier approaches were applied to HI. However, just a few studies have handled the issue up to now [135]. Recently, some DL-based techniques applied reconstruction error for recognition of outliers in other fields. However, they are not yet used in HI analysis.

- **Interpretable DL Model**

DL is usually disapproved of, since its decision-making process is not clear to individuals and thus frequently explained like a black box. People need to know the process of decision making or the basis of the decision. This might cause new findings in the domain of pathology. Even though this issue has not been fully resolved, some studies have tried to supply solutions, such as combined pathological pictures learning and diagnostic studies incorporated with interesting mechanisms [136]. In other fields, the basis of the decision might be ultimately displayed by visualizing the reaction of the deep network [137] or introducing a useful training picture applying impact functions [138].

- **Intraoperative Diagnosis**

Diagnosis by the pathologist during surgery impacts intraoperative decision making. Therefore, it might be another actual application in HI analysis. Because diagnosis time in an intraoperative examination is limited, a quick classifier while maintaining precision is significant. As a result of time limitation, the quick-freezing part is utilized rather than the formalin-fixed paraffin-embedded part that requires more time to get ready. Thus, for this reason, classification training must be executed, applying freezing part slides. Since the amount of appropriate WSI for analysis is not adequate, and function is more complicated than formalin-fixed paraffin-embedded slides, few studies have analyzed freezing parts [139].

- **Tumor-Infiltrating Immune Cell Analysis**

The microenvironment of carcinoma for immune cells has acquired significant interest recently. Thus, quantitative analysis for the carcinoma permeating of immune cells for slides applying ML methods is going to be among the emerging styles in HI analysis. Functions connected to analysis contain immune cell recognition in the H&E staining picture [140,141] and are recognized as a more specific form of immune cells applying immunohistochemistry [130]. Additionally, the structure of immune cell permeation and immune cell vicinity are supposedly linked to tumor treatment [142], spatial association analysis among immune cells and cells of cancer, and the association among this information and reaction to immunotherapy applying specific techniques, such as methods based on the graph [143], is likewise of good importance.

- **Challenges in HI analysis**

Typical DL architectures need their inputs in a particular structure with specific spatial dimensions. Moreover, these architectures are usually created for RGB pictures, while in digital HI, dealing with pictures in grayscale, HSV color might be desired for a particular system. Transforming pictures between color spaces, resizing pictures to suit GPU's storage, and determining the most effective resolution for applying at tiling are a few of the possible studies required, which will cause various levels of data loss. An acceptable information processing technique seeks to accomplish minimal data loss while using architectures for their maximal capacity. Input images are likely tiled or resized in most applications. It is also essential to balance the appropriate contexts and magnification with memory and computational constraints. Since CNN's can learn from smaller images more easily, images are not larger than the necessary context. A large amount of work was done to integrate low- and high-resolution inputs in different ways and issues to make better decisions [144].

- **Quality of training**

DL's accomplishment depends on the accessibility to high-quality training models to accomplish the required predictive efficiency [145,146]. Some efforts have been built to create extra annotated information by utilizing alternative methods, such as information augmentation [147], picture synthesis [104]. However, it is not even apparent that they are befitting from digital pathology.

- **Clinical translation**

There is a huge rapid development in AI research used in MI, and their possible effect has been shown by systems including the recognition of breast cancer metastasis [148], brain recognition [149], diagnosing diseases in retinal pictures [150], and so on. Regardless of this variety of systems, AI's actual and impactful implementation in medical practice will include many methods still to come.

- **Synthesis rather than marking**

An issue is that mapping of the label to the image domain is often unclear because the label mask can be mapped to many images. The training of the entire Generative Adversarial Network architecture can be difficult. The sizes of the regions-of-interest are given complexity here. Regions can display a diameter of up to several hundred pixels or thousands. This can be a big challenge, as the segmentation networks are implemented patch-wise.

- **Translation of morphology**

The optimal architecture for modified morphology settings does not show. Usually, unclear mappings can be particularly problematic in the event of morphological changes.

## 8. Conclusions

Different steps to analyze HIs are studied in this review for objective diagnosis automatically. In this survey, a comprehensive overview of different strategies in traditional and DL models has been presented. Different perspectives have tackled the analysis of HI for a wide variety of histology tasks (e.g., segmentation, tumor recognition, tissue classification). We have identified those that have been applied to various types of cancer (e.g., breast, kidney, colon, lung). For CAD in HI, there are primarily three phases: segmentation, feature extraction, and classification. The techniques developed for automatic analysis and evaluation of HIs help the pathologists in objective diagnosis for disease and decreased human error. A reference guide to recent literature methods for analyzing HIs manifests itself in the categorization techniques presented in this survey.

**Author Contributions:** Conceptualization, N.E. and M.E.; methodology, N.E., H.S., and M.E.; formal analysis, N.E. and M.E.; investigation, N.E., H.S., S.E.-S., S.M.R.I., and M.E.; writing—original draft preparation, N.E., H.S., and M.E.; writing—review and editing, N.E., H.S., S.E.-S., S.M.R.I., and M.E.; supervision, H.S. and M.E.; project administration, M.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.; van Ginneken, B.; Sanchez, I.C. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
2. Ker, J.; Wang, L.; Rao, J.; Lim, T. Deep Learning Applications in Medical Image Analysis. *IEEE Access* **2017**, *6*, 9375–9389. [[CrossRef](#)]
3. Perez, H.; Tah, J. Improving the Accuracy of Convolutional Neural Networks by Identifying and Removing Outlier Images in Datasets Using t-SNE. *Mathematics* **2020**, *8*, 662. [[CrossRef](#)]

4. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **2017**, *10*, 257–273. [[CrossRef](#)]
5. Pantanowitz, L. Digital images and the future of digital pathology. *J. Pathol. Inform.* **2010**, *1*. [[CrossRef](#)]
6. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological image analysis: A review. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 147–171. [[CrossRef](#)]
7. Greenspan, H.; Ginneken, B.; van Summers, R.M. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans Med. Imaging* **2016**, *35*, 1153–1159. [[CrossRef](#)]
8. Rubin, R.; Strayer, D.S.; Rubin, E. *Rubin's Pathology: Clinicopathologic Foundations of Medicine*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2008.
9. Hewitson, T.; Darby, I.; Walker, J. Histology Protocols. In *Methods in Molecular Biology*; Humana Press: Totowa, NJ, USA, 2010.
10. Li, C.; Chen, H.; Li, X.; Xu, N.; Hu, Z.; Xue, D.; Qi, S.; Ma, H.; Zhang, L.; Sun, H. A review for cervical histopathology image analysis using machine vision approaches. *Artif. Intell. Rev.* **2020**, *53*, 4821–4862. [[CrossRef](#)]
11. He, L.; Long, L.R.; Antani, S.; Thoma, G.R. Histology image analysis for carcinoma detection and grading. *Comput. Methods Programs Biomed.* **2012**, *107*, 538–556. [[CrossRef](#)]
12. Ghaznavi, F.; Evans, A.; Madabhushi, A.; Feldman, M. Digital Imaging in Pathology: Whole-Slide Imaging and Beyond. *Annu. Rev. Pathol. Mech. Dis.* **2013**, *8*, 331–359. [[CrossRef](#)]
13. Demir, C.; Yener, B. *Automated Cancer Diagnosis Based on Histopathological Images: A Systematic Survey*; Technical Report for Rensselaer Polytechnic Institute: New York, NY, USA, 2005.
14. Belsare, A. Histopathological Image Analysis Using Image Processing Techniques: An Overview. *Signal Image Process. Int. J.* **2012**, *3*, 23–36. [[CrossRef](#)]
15. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. Breast cancer histopathological image classification using Convolutional Neural Networks. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 January 2016; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2016; pp. 2560–2567.
16. Kieffer, B.; Babaie, M.; Kalra, S.; Tizhoosh, H.R. Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks. Proceeding of the 2017 Seventh International Conference on Image Processing Theory, Tools, and Applications (IPTA), Montreal, QC, Canada, 28 November–1 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
17. Mungle, T.; Tewary, S.; Das, D.; Arun, I.; Basak, B.; Agarwal, S.; Ahmed, R.; Chatterjee, S.; Chakraborty, C. MRF-ANN: A machine learning approach for automated ER scoring of breast cancer immunohistochemical images. *J. Microsc.* **2017**, *267*, 117–129. [[CrossRef](#)]
18. Sheikhzadeh, F.; Ward, R.K.; van Niekerk, D.; Guillaud, M. Automatic labeling of molecular biomarkers of immunohistochemistry images using fully convolutional networks. *PLoS ONE* **2018**, *13*, e0190783. [[CrossRef](#)]
19. Wang, D.; Foran, D.J.; Ren, J.; Zhong, H.; Kim, I.Y.; Qi, X. Exploring automatic prostate histopathology image gleason grading via local structure modeling. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2015; Volume 2015, pp. 2649–2652.
20. Roux, L.; Racoceanu, D.; Loménié, N.; Kulikova, M.; Irshad, H.; Klossa, J.; Capron, F.; Genestie, C.; Le Naour, G.; Gurcan, M.N. Mitosis detection in breast cancer histological images An ICPR 2012 contest. *J. Pathol. Inform.* **2013**, *4*, 8. [[CrossRef](#)]
21. Shah, M.; Wang, D.; Rubadue, C.; Suster, D.; Beck, A. Deep learning assessment of tumor proliferation in breast cancer histological images. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2017; pp. 600–603.
22. Chen, H.; Qi, X.; Yu, L.; Heng, P.-A. DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2016; pp. 2487–2496.

23. Gertych, A.; Ing, N.; Ma, Z.; Fuchs, T.J.; Salman, S.; Mohanty, S.; Bhele, S.; Velásquez-Vacca, A.; Amin, M.B.; Knudsen, B.S. Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Comput. Med. Imaging Graph.* **2015**, *46*, 197–208. [CrossRef]
24. Caicedo, J.C.; González, F.A.; Romero, E. Content-based histopathology image retrieval using a kernel-based semantic annotation framework. *J. Biomed. Inform.* **2011**, *44*, 519–528. [CrossRef]
25. Caie, P.D.; Turnbull, A.K.; Farrington, S.M.; Oniscu, A.; Harrison, D.J. Quantification of tumour budding, lymphatic vessel density and invasion through image analysis in colorectal cancer. *J. Transl. Med.* **2014**, *12*, 156. [CrossRef]
26. Sirinukunwattana, K.; Pluim, J.P.; Chen, H.; Qi, X.; Heng, P.-A.; Guo, Y.B.; Wang, L.Y.; Matuszewski, B.J.; Bruni, E.; Sanchez, U.; et al. Gland segmentation in colon histology images: The glas challenge contest. *Med. Image Anal.* **2017**, *35*, 489–502. [CrossRef]
27. Qi, X.; Wang, D.; Rodero, I.; Diaz-Montes, J.; Gensure, R.H.; Xing, F.; Zhong, H.; Goodell, L.; Parashar, M.; Foran, D.J.; et al. Content-based histopathology image retrieval using CometCloud. *BMC Bioinform.* **2014**, *15*, 1–17. [CrossRef]
28. Sparks, R.; Madabhushi, A. Out-of-Sample Extrapolation utilizing Semi-Supervised Manifold Learning (OSE-SSL): Content Based Image Retrieval for Histopathology Images. *Sci. Rep.* **2016**, *6*, 27306. [CrossRef]
29. Sridhar, A.; Doyle, S.; Madabhushi, A. Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces. *J. Pathol. Inform.* **2015**, *6*. [CrossRef]
30. Vanegas, J.A.; Arevalo, J.; González, F.A. Unsupervised feature learning for content-based histopathology image retrieval. In *2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI), Klagenfurt, Austria, 18–20 June 2014*; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2014; pp. 1–6.
31. Zhang, X.; Liu, W.; Dundar, M.; Badve, S.; Zhang, S. Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval. *IEEE Trans. Med. Imaging* **2014**, *34*, 496–506. [CrossRef]
32. Flahou, B.; Haesebrouck, F.; Smet, A. Non-Helicobacter pylori Helicobacter Infections in Humans and Animals. In *Helicobacter Pylori Research*; Springer: Tokyo, Japan, 2016; pp. 233–269. [CrossRef]
33. Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.M.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **2013**, *45*, 1113–1120. [CrossRef]
34. Molin, M.D.; Matthaei, H.; Wu, J.; Blackford, A.; Debeljak, M.; Rezaee, N.; Wolfgang, C.L.; Butturini, G.; Salvia, R.; Bassi, C.; et al. Clinicopathological Correlates of Activating GNAS Mutations in Intraductal Papillary Mucinous Neoplasm (IPMN) of the Pancreas. *Ann. Surg. Oncol.* **2013**, *20*, 3802–3808. [CrossRef]
35. Yoshida, A.; Tsuta, K.; Nakamura, H.; Kohno, T.; Takahashi, F.; Asamura, H.; Sekine, I.; Fukayama, M.; Shibata, T.; Furuta, K.; et al. Comprehensive Histologic Analysis of ALK-Rearranged Lung Carcinomas. *Am. J. Surg. Pathol.* **2011**, *35*, 1226–1234. [CrossRef]
36. Arevalo, J.; Cruz-Roa, A. Histopathology image representation for automatic analysis: A state-of-the-art review. *Rev. Med.* **2014**, *22*, 79–91. [CrossRef]
37. Lyon, H.O.; de Leenheer, A.P.; Horobin, R.W.; Lambert, W.E.; Schulte, E.K.W.; van Liedekerke, B.; Wittekind, D.H. Standardization of reagents and methods used in cytological and histological practice with emphasis on dyes, stains and chromogenic reagents. *J. Mol. Histol.* **1994**, *26*, 533–544. [CrossRef]
38. Khan, A.M.; 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*, 1729–1738. [CrossRef]
39. Anghel, A.; Stanisavljevic, M.; Andani, S.; Papandreou, N.; Rüschoff, J.H.; Wild, P.; Gabrani, M.; Pozidis, H. A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology. *Front. Med.* **2019**, *6*. [CrossRef]
40. Can, A.; Bello, M.; Cline, H.E.; Tao, X.; Ginty, F.; Sood, A.; Gerdes, M.; Montalto, M. Multi-modal imaging of histological tissue sections. In Proceedings of the 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, France, 14–18 May 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 288–291. [CrossRef]
41. Casiraghi, E.; Cossa, M.; Huber, V.; Tozzi, M.; Rivoltini, L.; Villa, A.; Vergani, B. MIAQuant, a novel system for automatic segmentation, measurement, and localization comparison of different biomarkers from serialized histological slices. *Eur. J. Histochem.* **2017**, *61*, 61. [CrossRef]

42. Casiraghi, E.; Huber, V.; Frasca, M.; Cossa, M.; Tozzi, M.; Rivoltini, L.; Leone, B.E.; Villa, A.; Vergani, B. A novel computational method for automatic segmentation, quantification and comparative analysis of immunohistochemically labeled tissue sections. *BMC Bioinform.* **2018**, *19*, 357–397. [[CrossRef](#)]
43. Roy, S.; Jain, A.K.; Lal, S.; Kini, J. A study about color normalization methods for histopathology images. *Micron* **2018**, *114*, 42–61. [[CrossRef](#)]
44. Mărginean, R.; Andreica, A.; Dioşan, L.; Bálint, Z. Feasibility of Automatic Seed Generation Applied to Cardiac MRI Image Analysis. *Mathematics* **2020**, *8*, 1511. [[CrossRef](#)]
45. Gleason, D.F. Histologic grading of prostate cancer: A perspective. *Hum. Pathol.* **1992**, *23*, 273–279. [[CrossRef](#)]
46. Kong, J.; Sertel, O.; Shimada, H.; Boyer, K.L.; Saltz, J.H.; Gurcan, M.N. Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation. *Pattern Recognit.* **2009**, *42*, 1080–1092. [[CrossRef](#)]
47. Washington, K.; Berlin, J.; Branton, P.; Burgart, L.J.; Carter, D.K.; Fitzgibbons, P.L.; Halling, K.; Frankel, W.; Jessup, J.; Kakar, S.; et al. Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum. *Arch. Pathol. Lab. Med.* **2009**, *133*, 1539–1551. [[CrossRef](#)]
48. Irshad, H.; Veillard, A.; Roux, L.; Racoceanu, D. Methods for Nuclei Detection, Segmentation, and Classification in Digital Histopathology: A Review—Current Status and Future Potential. *IEEE Rev. Biomed. Eng.* **2013**, *7*, 97–114. [[CrossRef](#)]
49. Dalle, J.-R.; Li, H.; Huang, C.-H.; Leow, W.K.; Racoceanu, D.; Putti, T. Nuclear pleomorphism scoring by selective cell nuclei detection. In Proceedings of the Workshop on Applications of Computer Vision, Snowbird, UT, USA, 7–8 December 2009.
50. Wahlby, C.; Sintorn, I.-M.; Erlandsson, F.; Borgefors, G.; Bengtsson, E. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *J. Microsc.* **2004**, *215*, 67–76. [[CrossRef](#)]
51. Jung, C.; Kim, C. Segmenting Clustered Nuclei Using H-minima Transform-Based Marker Extraction and Contour Parameterization. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 2600–2604. [[CrossRef](#)]
52. Cosatto, E.; Miller, M.; Graf, H.P.; Meyer, J.S. Grading Nuclear Pleomorphism on Histological Micrographs. In Proceedings of the 2008 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2008; pp. 1–4.
53. Al-Kofahi, Y.; Lassoued, W.; Lee, W.; Roysam, B. Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology Images. *IEEE Trans. Biomed. Eng.* **2009**, *57*, 841–852. [[CrossRef](#)]
54. Veta, M.; Huisman, A.; Viergever, M.; van Diest, P.J.; Pluim, J. Marker-controlled watershed segmentation of nuclei in H&E stained breast cancer biopsy images. In Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, 30 March–2 April 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 618–621. [[CrossRef](#)]
55. Aptoula, E.; Courty, N.; Lefèvre, S. Mitosis detection in breast cancer histological images with mathematical morphology. In Proceedings of the 2013 21st Signal Processing and Communications Applications Conference (SIU), Haspolat, Turkey, 24–26 April 2013; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2013; pp. 1–4.
56. Ciresan, D.C.; Giusti, A.; Gambardella, L.M.; Schmidhuber, J. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, Japan, 22–26 September 2013; Springer: Berlin/Heidelberg, Germany, 2013; Volume 2, pp. 411–418.
57. Petushi, S.; Garcia, F.U.; Haber, M.M.; Katsinis, C.; Tozeren, A. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC Med Imaging* **2006**, *6*, 14. [[CrossRef](#)]
58. Rittscher, J.; Machiraju, R.; Wong, S. *Microscopic Image Analysis for Life Science Applications*; Artech House: Norwood, MA, USA, 2008.
59. Boucheron, L.E. *Object- and Spatial-Level Quantitative Analysis of Multispectral Histopathology Images for Detection and Characterization of Cancer*; University of California at Santa Barbara: Santa Barbara, CA, USA, 2008.
60. Kuse, M.; Sharma, T.; Gupta, S. A Classification Scheme for Lymphocyte Segmentation in H&E Stained Histology Images. In *Static Analysis*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2010; pp. 235–243.

61. Chekkoury, A.; Khurd, P.; Ni, J.; Bahlmann, C.; Kamen, A.; Patel, A.; Grady, L.; Singh, M.; Groher, M.; Navab, N.; et al. Automated malignancy detection in breast histopathological images. In *Medical Imaging 2012: Computer-Aided Diagnosis*; SPIE: Bellingham, WA, USA, 2012; pp. 332–344.
62. Doyle, S.; Hwang, M.; Shah, K.; Madabhushi, A.; Feldman, M.D.; Tomaszewski, J.E. Automated Grading of Prostate Cancer Using Architectural and Textural Image Features. In Proceedings of the 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Arlington, VA, USA, 12–15 April 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1284–1287.
63. Di Franco, M.D.; O’Hurley, G.; Kay, E.W.; Watson, R.W.G.; Cunningham, P. Ensemble-based system for whole-slide prostate cancer probability mapping using color texture features. *Comput. Med. Imaging Graphics* **2011**, *35*, 629–645. [[CrossRef](#)]
64. Huang, P.-W.; Lai, Y.H. Effective segmentation and classification for HCC biopsy images. *Pattern Recognit.* **2010**, *43*, 1550–1563. [[CrossRef](#)]
65. Alexandratou, E.; Atlamazoglou, V.; Thireou, T.; Agrogiannis, G.; Togas, D.; Kavantzas, N.; Patsouris, E.; Yova, D. Evaluation of machine learning techniques for prostate cancer diagnosis and Gleason grading. *Int. J. Comput. Intell. Bioinform. Syst. Biol.* **2010**, *1*, 297. [[CrossRef](#)]
66. Basavanhally, A.; Yu, E.; Xu, J.; Ganesan, S.; Feldman, M.; Tomaszewski, J.E.; Madabhushi, A. Incorporating domain knowledge for tubule detection in breast histopathology using O’Callaghan neighborhoods. *Inter. Soc. Optics Photonics* **2011**, 7963, 796310.
67. Al-Kadi, O.S. Texture measures combination for improved meningioma classification of histopathological images. *Pattern Recognit.* **2010**, *43*, 2043–2053. [[CrossRef](#)]
68. Demir, C.; Kandemir, M.; Tosun, A.B.; Sokmensuer, C. Automatic segmentation of colon glands using object-graphs. *Med. Image Anal.* **2010**, *14*, 1–12. [[CrossRef](#)]
69. Tosun, A.B.; Demir, C. Graph Run-Length Matrices for Histopathological Image Segmentation. *IEEE Transact. Med. Imaging* **2011**, *30*, 721–732. [[CrossRef](#)]
70. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
71. Le Cun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
72. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
73. Arevalo, J.; Cruz-Roa, A.; González, F.A. Hybrid image representation learning model with invariant features for basal cell carcinoma detection. In Proceedings of the IX International Seminar on Medical Information Processing and Analysis, Mexico City, Mexico, 11–14 November 2013; SPIE: Bellingham, WA, USA, 2013; Volume 8922, p. 89220M.
74. Nayak, N.; Chang, H.; Borowsky, A.; Spellman, P.T.; Parvin, B. Classification of tumor histopathology via sparse feature learning. In Proceedings of the 2013 IEEE 10th International Symposium on Biomedical Imaging, San Francisco, CA, USA, 7–11 April 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 410–413.
75. Malon, C.D.; Cosatto, E. Classification of mitotic figures with convolutional neural networks and seeded blob features. *J. Pathol. Informat.* **2013**, *4*, 9. [[CrossRef](#)]
76. Xu, Y.; Mo, T.; Feng, Q.; Zhong, P.; Lai, M.; Chang, E.I.-C. Deep learning of feature representation with multiple instance learning for medical image analysis. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2014; pp. 1626–1630.
77. Hou, L.; Samaras, D.; Kurç, T.M.; Gao, Y.; Davis, J.E.; Saltz, J.H. Efficient Multiple Instance Convolutional Neural Networks for Gigapixel Resolution Image Classification. *arXiv* **2015**, arXiv:abs/1504.07947.
78. Arevalo, J.; Cruz-Roa, A.; Arias, V.; Romero, E.; González, F.A. An unsupervised feature learning framework for basal cell carcinoma image analysis. *Artif. Intell. Med.* **2015**, *64*, 131–145. [[CrossRef](#)]
79. Chang, H.; Zhou, Y.; Borowsky, A.; Barner, K.; Spellman, P.; Parvin, B. Stacked Predictive Sparse Decomposition for Classification of Histology Sections. *Int. J. Comput. Vis.* **2014**, *113*, 3–18. [[CrossRef](#)] [[PubMed](#)]

80. Han, J.; Fontenay, G.V.; Wang, Y.; Mao, J.-H.; Chang, H. Phenotypic characterization of breast invasive carcinoma via transferable tissue morphometric patterns learned from glioblastoma multiforme. Proceeding of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1025–1028.
81. Noël, H.; Roux, L.; Lu, S.; Boudier, T. Detection of high-grade atypia nuclei in breast cancer imaging. In *Medical Imaging*; SPIE: Bellingham, WA, USA, 2015; p. 94200R.
82. Romo-Bucheli, D.; Janowczyk, A.; Gilmore, H.; Romero, E.; Madabhushi, A. Automated Tubule Nuclei Quantification and Correlation with Oncotype DX risk categories in ER+ Breast Cancer Whole Slide Images. *Sci. Rep.* **2016**, *6*, 32706. [CrossRef] [PubMed]
83. Chen, T. Deep Learning Based Automatic Immune Cell Detection for Immunohistochemistry Images. In *International Workshop on Machine Learning in Medical Imaging*; Springer: Cham, Switzerland, 2014; pp. 17–24.
84. Srinidhi, C.L.; Ciga, O.; Martel, A.L. Deep neural network models for computational histopathology: A survey. *Med Image Anal.* **2020**, *101813*. [CrossRef] [PubMed]
85. Sumi, P.S.; Delhibabu, R. Glioblastoma Multiforme Classification On High Resolution Histology Image Using Deep Spatial Fusion Network. In Proceedings of theCEUR Workshop, Como, Italy, 9–11 September 2019.
86. Zhang, L.; Wu, Y.; Zheng, B.; Su, L.; Chen, Y.; Ma, S.; Hu, Q.; Zou, X.; Yao, L.; Yang, Y.; et al. Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy. *Theranostics* **2019**, *9*, 2541–2554. [CrossRef] [PubMed]
87. Agarwal, A. GPU Based Digital Histopathology and Diagnostic Support System for Breast Cancer Detection: A Comparison of CNN Models and Machine Learning Models. *Nature Rev. Drug Discov.* **2019**, *18*, 463–477.
88. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
89. Tschuchnig, M.E.; Oostingh, G.J.; Gadermayr, M. Generative Adversarial Networks in Digital Pathology: A Survey on Trends and Future Potential. *arXiv* **2020**, arXiv:abs/2004.14936. [CrossRef]
90. Litjens, G.J.S.; Sánchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.D.; Kovacs, I.; Kaa, C.H.; van de Bult, P.; Ginneken, B.; van Laak, J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Reports* **2016**, *6*. [CrossRef]
91. Nagpal, K.; Foote, D.; Liu, Y.; Chen, P.-H.C.; Wulczyn, E.; Tan, F.; Olson, N.; Smith, J.L.; Mohtashamian, A.; Wren, J.H.; et al. Publisher Correction: Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *npj Digit. Med.* **2019**, *2*, 2. [CrossRef] [PubMed]
92. Zhao, Z.; Lin, H.; Chen, H.; Heng, P.-A. PFA-ScanNet: Pyramidal Feature Aggregation with Synergistic Learning for Breast Cancer Metastasis Analysis. *Lecture Notes Comput. Sci.* **2019**, 586–594.
93. Xing, F.; Xie, Y.; Yang, L. An Automatic Learning-Based Framework for Robust Nucleus Segmentation. *IEEE Trans. Med. Imaging* **2015**, *35*, 550–566. [CrossRef]
94. Gu, F.; Burlutskiy, N.; Andersson, M.; Wilén, L.K. Multi-resolution Networks for Semantic Segmentation in Whole Slide Images. In *Lecture Notes in Computer Science*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2018; pp. 11–18.
95. Tellez, D.; Balkenhol, M.; Otte-Holler, I.; van de Loo, R.; Vogels, R.; Bult, P.; Wauters, C.; Vreuls, W.; Mol, S.; Karssemeijer, N.; et al. Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks. *IEEE Trans. Med. Imaging* **2018**, *37*, 2126–2136. [CrossRef]
96. Wei, J.W.; Tafe, L.J.; Linnik, Y.A.; Vaickus, L.J.; Tomita, N.; Hassanpour, S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Sci. Rep.* **2019**, *9*, 1–8. [CrossRef]
97. Song, Y.; Tan, E.-L.; Jiang, X.; Cheng, J.-Z.; Ni, D.; Chen, S.; Lei, B.Y.; Wang, T. Accurate Cervical Cell Segmentation from Overlapping Clumps in Pap Smear Images. *IEEE Transact. Med. Imaging* **2017**, *36*, 288–300. [CrossRef]
98. Agarwalla, A.; Shaban, M.; Rajpoot, N.M. Representation-Aggregation Networks for Segmentation of Multi-Gigapixel Histology Images. *arXiv* **2017**, arXiv:1707-08814.
99. Ding, H.; Pan, Z.; Cen, Q.; Li, Y.; Chen, S. Multi-scale fully convolutional network for gland segmentation using three-class classification. *Neurocomputing* **2020**, *380*, 150–161. [CrossRef]

100. Bejnordi, B.E.; Zuidhof, G.C.A.; Balkenhol, M.; Hermsen, M.; Bult, P.; Ginneken, B.; van Karssemeijer, N.; Litjens, G.J.S.; Laak, J. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J. Med. Imaging* **2017**, *4*, 044504. [[CrossRef](#)]
101. Seth, N.; Akbar, S.; Nofech-Mozes, S.; Salama, S.; Martel, A.L. Automated Segmentation of DCIS in Whole Slide Images. In *Case-Based Reasoning Research and Development*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; pp. 67–74.
102. Xu, J.; Xiang, L.; Liu, Q.; Gilmore, H.; Wu, J.; Tang, J.; Madabhushi, A. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Trans. Med Imaging* **2015**, *35*, 119–130. [[CrossRef](#)]
103. Bulten, W.; Litjens, G. Unsupervised Prostate Cancer Detection on H&E using Convolutional Adversarial Autoencoders. *arXiv* **2018**, arXiv:1804.07098.
104. Hou, L.; Agarwal, A.; Samaras, D.; Kurc, T.M.; Gupta, R.R.; Saltz, J.H. Robust Histopathology Image Analysis: To Label or to Synthesize? In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2019; pp. 8525–8534.
105. Sari, C.T.; Gunduz-Demir, C. Unsupervised Feature Extraction via Deep Learning for Histopathological Classification of Colon Tissue Images. *IEEE Transact. Med. Imaging* **2019**, *38*, 1139–1149. [[CrossRef](#)] [[PubMed](#)]
106. Gadermayr, M.; Gupta, L.; Appel, V.; Boor, P.; Klinkhammer, B.M.; Merhof, D. Generative Adversarial Networks for Facilitating Stain-Independent Supervised and Unsupervised Segmentation: A Study on Kidney Histology. *IEEE Trans. Med Imaging* **2019**, *38*, 2293–2302. [[CrossRef](#)] [[PubMed](#)]
107. Gadermayr, M.; Gupta, L.; Klinkhammer, B.M.; Boor, P.; Merhof, D. Unsupervisedly Training GANs for Segmenting Digital Pathology with Automatically Generated Annotations. *arXiv* **2018**, arXiv:1805.10059.
108. Komura, D.; Ishikawa, S. Machine Learning Methods for Histopathological Image Analysis. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 34–42. [[CrossRef](#)]
109. Search Home—Biospecimen Research Database. Available online: <https://brd.nci.nih.gov/brd/image-search/searchhome> (accessed on 1 October 2020).
110. TMAD Main Menu. Available online: <https://tma.im/cgi-bin/home.pl> (accessed on 1 October 2020).
111. Home—CAMELYON17—Grand Challenge. Available online: <https://camelyon17.grand-challenge.org/> (accessed on 1 October 2020).
112. Breast Cancer Histopathological Database (BreakHis)—Laboratório Visão Robótica e Imagem. Available online: <https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/> (accessed on 1 October 2020).
113. Search GDC. Available online: <https://portal.gdc.cancer.gov/legacy-archive/search/f> (accessed on 1 October 2020).
114. Marinelli, R.J.; Montgomery, K.; Liu, C.L.; Shah, N.H.; Prapong, W.; Nitzberg, M.; Zachariah, Z.K.; Sherlock, G.; Natkunam, Y.; West, R.B.; et al. The Stanford Tissue Microarray Database. *Nucleic Acids Res.* **2007**, *36*, D871–D877. [[CrossRef](#)]
115. Dataset Tumor Proliferation Assessment Challenge 2016. Available online: <http://tupac.tue-image.nl/node/3> (accessed on 1 October 2020).
116. Bentaieb, A.; Li-Chang, H.; Huntsman, D.; Hamarneh, G. A structured latent model for ovarian carcinoma subtyping from histopathology slides. *Med Image Anal.* **2017**, *39*, 194–205. [[CrossRef](#)]
117. Ovarian Carcinomas Histopathology Dataset. Available online: <http://ensc-mica-www02.ensc.sfu.ca/download/> (accessed on 1 October 2020).
118. Babaie, M.; Kalra, S.; Sriram, A.; Mitcheltree, C.; Zhu, S.; Khatami, A.; Rahnamayan, S.; Tizhoosh, H.R. Classification and Retrieval of Digital Pathology Scans: A New Dataset. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–16 July 2017; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2017; pp. 760–768.
119. Kumar, N.; Verma, R.; Sharma, S.; Bhargava, S.; Vahadane, A.; Sethi, A. A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology. *IEEE Trans. Med Imaging* **2017**, *36*, 1550–1560. [[CrossRef](#)]
120. Pathology Images: KIMIA Path960—Kimia Lab. Available online: <https://kimialab.uwaterloo.ca/kimia/index.php/pathology-images-kimia-path960/> (accessed on 1 October 2020).

121. Gelasca, E.D.; Byun, J.; Obara, B.; Manjunath, B.S. Evaluation and benchmark for biological image segmentation. In Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2008; pp. 1816–1819.
122. Bio-Segmentation Center for Bio-Image Informatics UC Santa Barbara. Available online: <https://bioimage.ucsb.edu/research/bio-segmentation> (accessed on 1 October 2020).
123. Bioimaging Challenge 2015 Breast Histology Dataset—Datasets CKAN. Available online: <https://rdm.inesctec.pt/dataset/nis-2017-003> (accessed on 1 October 2020).
124. BIALab@Warwick: GlaS Challenge Contest. Available online: <https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/> (accessed on 1 October 2020).
125. Kather, J.N.; Marx, A.; Reyes-Aldasoro, C.C.; Schad, L.R.; Zoellner, F.G.; Weis, C.-A. Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. *Oncotarget* **2015**, *6*, 19163–19176. [CrossRef]
126. Dataset—MITOS-ATYPIA-14—Grand Challenge. Available online: <https://mitos-atypia-14.grand-challenge.org/dataset/> (accessed on 1 October 2020).
127. Nucleisegmentation. Available online: <https://nucleisegmentationbenchmark.weebly.com/> (accessed on 1 October 2020).
128. Janowczyk, A.; Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Pathol. Informat.* **2016**, *7*. [CrossRef]
129. Andrew Janowczyk—Tidbits from Along the Way. Available online: <http://www.andrewjanowczyk.com/> (accessed on 1 October 2020).
130. Ma, Z.; Shiao, S.L.; Yoshida, E.J.; Swartwood, S.; Huang, F.; Doche, M.E.; Chung, A.P.; Knudsen, B.S.; Gertych, A. Data integration from pathology slides for quantitative imaging of multiple cell types within the tumor immune cell infiltrate. *Diagn. Pathol.* **2017**, *12*. [CrossRef] [PubMed]
131. Linder, N.; Konsti, J.; Turkki, R.; Rahtu, E.; Lundin, M.; Nordling, S.; Haglund, C.; Ahonen, T.; Pietikäinen, M.; Lundin, J. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. *Diagn. Pathol.* **2012**, *7*, 22. [CrossRef] [PubMed]
132. Egfr Colon Stroma Classification. Available online: <http://fimm.webmicroscope.net/supplements/epistroma> (accessed on 1 October 2020).
133. Jimenez-del-Toro, O.; Otálora, S.; Andersson, M.; Eurén, K.; Hedlund, M.; Rousson, M.; Müller, H.; Atzori, M. Chapter 10—Analysis of Histopathology Images: From Traditional Machine Learning to Deep Learning; Academic Press: Cambridge, MA, USA, 2017; pp. 281–314.
134. Zhang, Y.; Zhang, B.; Coenen, F.; Xiao, J.; Lu, W. One-class kernel subspace ensemble for medical image classification. *EURASIP J. Adv. Signal Process.* **2014**, *2014*, 17. [CrossRef]
135. Xia, Y.; Cao, X.; Wen, F.; Hua, G.; Sun, J. Learning Discriminative Reconstructions for Unsupervised Outlier Removal. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Las Condes, Chile, 11–15 December 2015; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2015; pp. 1511–1519.
136. Samek, W.; Binder, A.; Montavon, G.; Lapuschkin, S.; Muller, K.-R. Evaluating the Visualization of What a Deep Neural Network Has Learned. *IEEE Trans. Neural Networks Learn. Syst.* **2016**, *28*, 2660–2673. [CrossRef] [PubMed]
137. Zintgraf, L.M.; Cohen, T.S.; Adel, T.; Welling, M. Visualizing Deep Neural Network Decisions: Prediction Difference Analysis. *arXiv* **2017**, arXiv:1702.04595.
138. Koh, P.W.; Liang, P. Understanding Black-box Predictions via Influence Functions. *arXiv* **2017**, arXiv:abs/1703.04730.
139. Abas, F.S.; Gokozan, H.; Goksel, B.; Otero, J.J. *Intraoperative Neuropathology of Glioma Recurrence: Cell Detection and Classification*; SPIE: Bellingham, WA, USA, 2016; p. 979109. [CrossRef]
140. Chen, J.; Srinivas, C. Automatic Lymphocyte Detection in H&E Images with Deep Neural Networks. *arXiv* **2016**, arXiv:abs/1612.03217.
141. Turkki, R.; Linder, N.; Kovanen, P.E.; Pellinen, T.; Lundin, J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *J. Pathol. Informat.* **2016**, *7*, 38. [CrossRef]

142. Feng, Z.; Bethmann, D.; Kappler, M.; Ballesteros-Merino, C.; Eckert, A.; Bell, R.B.; Cheng, A.; Bui, T.; Leidner, R.; Urba, W.J.; et al. Multiparametric immune profiling in HPV—Oral squamous cell cancer. *JCI Insight* **2017**, *2*, e93652. [[CrossRef](#)]
143. Basavanhally, A.N.; Ganesan, S.; Agner, S.; Monaco, J.P.; Feldman, M.D.; Tomaszewski, J.E.; Bhanot, G.; Madabhushi, A. Computerized Image-Based Detection and Grading of Lymphocytic Infiltration in HER2+ Breast Cancer Histopathology. *IEEE Trans. Biomed. Eng.* **2009**, *57*, 642–653. [[CrossRef](#)]
144. Li, J.; Li, W.; Gertych, A.; Knudsen, B.S.; Speier, W.; Arnold, C. An attention-based multi-resolution model for prostate whole slide imageclassification and localization. *arXiv* **2019**, arXiv:1905.13208.
145. Bera, K.; Schalper, K.A.; Rimm, D.L.; Velcheti, V.; Madabhushi, A. Artificial intelligence in digital pathology new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 703–715. [[CrossRef](#)] [[PubMed](#)]
146. Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital pathology and artificial intelligence. *Lancet Oncol.* **2019**, *20*, 253–261. [[CrossRef](#)]
147. Tellez, D.; Litjens, G.J.S.; Bárdi, P.; Bulten, W.; Bokhorst, J.-M.; Ciompi, F.; Laak, J. Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Med. Image Anal.* **2019**, *58*, 101544. [[CrossRef](#)] [[PubMed](#)]
148. Steiner, D.F.; Macdonald, R.; Liu, Y.; Truszkowski, P.; Hipp, J.D.; Gammage, C.; Thng, F.; Peng, L.; Stumpe, M.C. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am. J. Surg. Pathol.* **2018**, *42*, 1636–1646. [[CrossRef](#)]
149. Kamnitsas, K.; Ferrante, E.; Parisot, S.; Ledig, C.; Nori, A.V.; Criminisi, A.; Rueckert, D.; Glocker, B. *DeepMedic for Brain Tumor Segmentation*; Springer: Berlin/Heidelberg, Germany, 2016.
150. Voets, M.; Møllersen, K.; Bongo, L.A. Replication study: Development and validation of deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *arXiv* **2018**, arXiv:abs/1803.04337.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).