

Review

A Review of Computer-Aided Expert Systems for Breast Cancer Diagnosis

Xin Yu Liew *, Nazia Hameed and Jeremie Clos 

Jubilee Campus, University of Nottingham, Wollaton Road, Nottingham NG8 1BB, UK;
nazia.hameed@nottingham.ac.uk (N.H.); jeremie.clos@nottingham.ac.uk (J.C.)

* Correspondence: hcyx11@nottingham.ac.uk

Simple Summary: Breast cancer is one of the most commonly diagnosed diseases in females around the world. The most threatening is when cancer spreads uncontrollably to other parts of the body and can cause death. Early detection of breast cancer lowers the risk of death among patients and enables appropriate treatments to control the progression of cancer. To diagnose breast cancer, high complex visuals of the breast tissue can be collected through histopathology images that provide informative details which validate the stage of the cancer. The aim of this study is to investigate techniques applied in histopathology images in diagnosing breast cancer.

Abstract: A computer-aided diagnosis (CAD) expert system is a powerful tool to efficiently assist a pathologist in achieving an early diagnosis of breast cancer. This process identifies the presence of cancer in breast tissue samples and the distinct type of cancer stages. In a standard CAD system, the main process involves image pre-processing, segmentation, feature extraction, feature selection, classification, and performance evaluation. In this review paper, we reviewed the existing state-of-the-art machine learning approaches applied at each stage involving conventional methods and deep learning methods, the comparisons within methods, and we provide technical details with advantages and disadvantages. The aims are to investigate the impact of CAD systems using histopathology images, investigate deep learning methods that outperform conventional methods, and provide a summary for future researchers to analyse and improve the existing techniques used. Lastly, we will discuss the research gaps of existing machine learning approaches for implementation and propose future direction guidelines for upcoming researchers.

Keywords: machine learning; deep learning; computer aided diagnosis; breast cancer; histopathology images; classification; medical imaging



Citation: Liew, X.Y.; Hameed, N.; Clos, J. A Review of Computer-Aided Expert Systems for Breast Cancer Diagnosis. *Cancers* **2021**, *13*, 2764. <https://doi.org/10.3390/cancers13112764>

Academic Editor: Ognjen Arandjelović

Received: 26 April 2021

Accepted: 28 May 2021

Published: 2 June 2021

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



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The human body is formed of trillions of cells. ‘Cancer’ is a term used when a cell divides abnormally or uncontrollably, which can happen in various parts of the body. The disease type is categorised based on which part of the body cancer occurs. This situation, if left unchecked, will lead to death. Amongst the distinct types of cancer, the most common type of cancer happening in females is breast cancer. According to the World Health Organisation (WHO), breast cancer is the most frequent cancer among women, affecting 2.1 million women each year. About 627,000 women died from breast cancer in 2018, which accounted for around 15% of all cancer deaths among women [1]. In the United Kingdom, there are around 55,200 newly diagnosed breast cancer cases every year, which is about 150 every day from 2015 to 2017 [2]. From Figure 1, we can observe that breast cancer has the highest number of diagnosis incidence among all the common cancers [1].

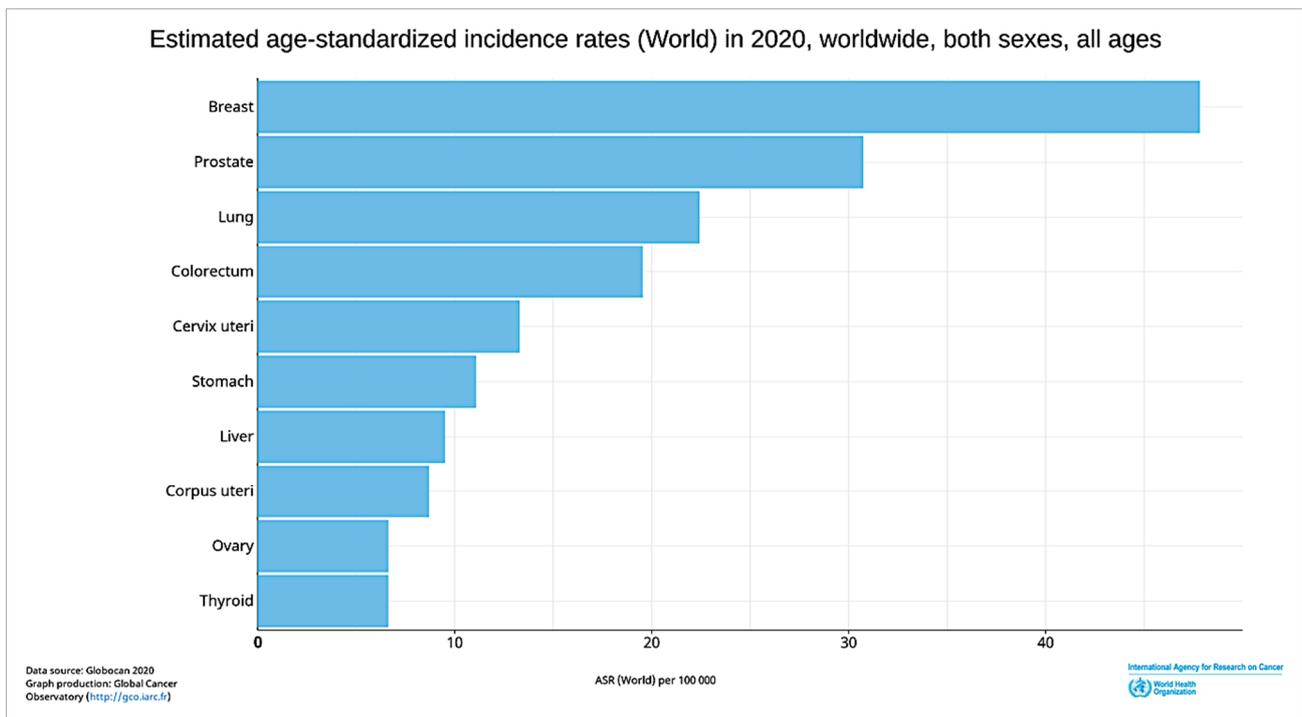


Figure 1. Incidence rates of cancer in 2020 [1].

However, it has been proven that an early detection of breast cancer can significantly increase the chances of successful treatment plan and ensure a long-term survival of the patients [3]. Statistically, if the disease is detected and diagnosed at an early stage, nearly all (98%) patients will survive for five years or more, compared to around 1 in 4 (26%) people when the disease is diagnosed at a later stage [2]. According to the most common procedure, a ‘two-week wait’ is the procedure to diagnose breast cancer [2]. The standard procedure to diagnose breast cancer by pathologists usually requires extensive microscopic assessment. Therefore, having an automated solution like a computer-aided diagnosis (CAD) system not only contributes to an easier diagnostic process, but also reduces the subjectivity in diagnosis.

With the advanced development of artificial intelligence, many machine learning techniques have been applied for CAD systems. This technique can potentially outperform humans and learn more efficiently with time, therefore integrating machine learning in diagnosis can supply useful knowledge to assist pathologists in evaluating and analysing enormous amounts of medical data [4]. It could also speed up the process due to the capability to process large data much faster than manual diagnosis by a pathologist [4]. Breast cancer diagnosis can be considered as a classification problem in machine learning, in which the result indicates which class of cancer it belongs to. Fundamentally, the main steps involved in developing the core of a computer-aided diagnosis (CAD) system for breast cancer are presented in Figure 2.

Conventionally, several popular machine learning algorithms applied to classification problems include naïve Bayes [5], artificial neural network [6], support vector machine (SVM) [7], and many more. However, these algorithms might not have the ability to consider issues such as imbalanced costs of misclassification within various classes, leading to undesired consequences. Recently, deep learning methods were introduced to improve on conventional machine learning methods by extracting information automatically as part of the learning process, leading to undoubtedly better solutions [8]. Deep learning was shown to outperform state-of-the-art methods in many fields of medical imaging analysis tasks. Therefore, in this paper we will discuss and compare both approaches applied to develop a CAD system for the breast cancer.

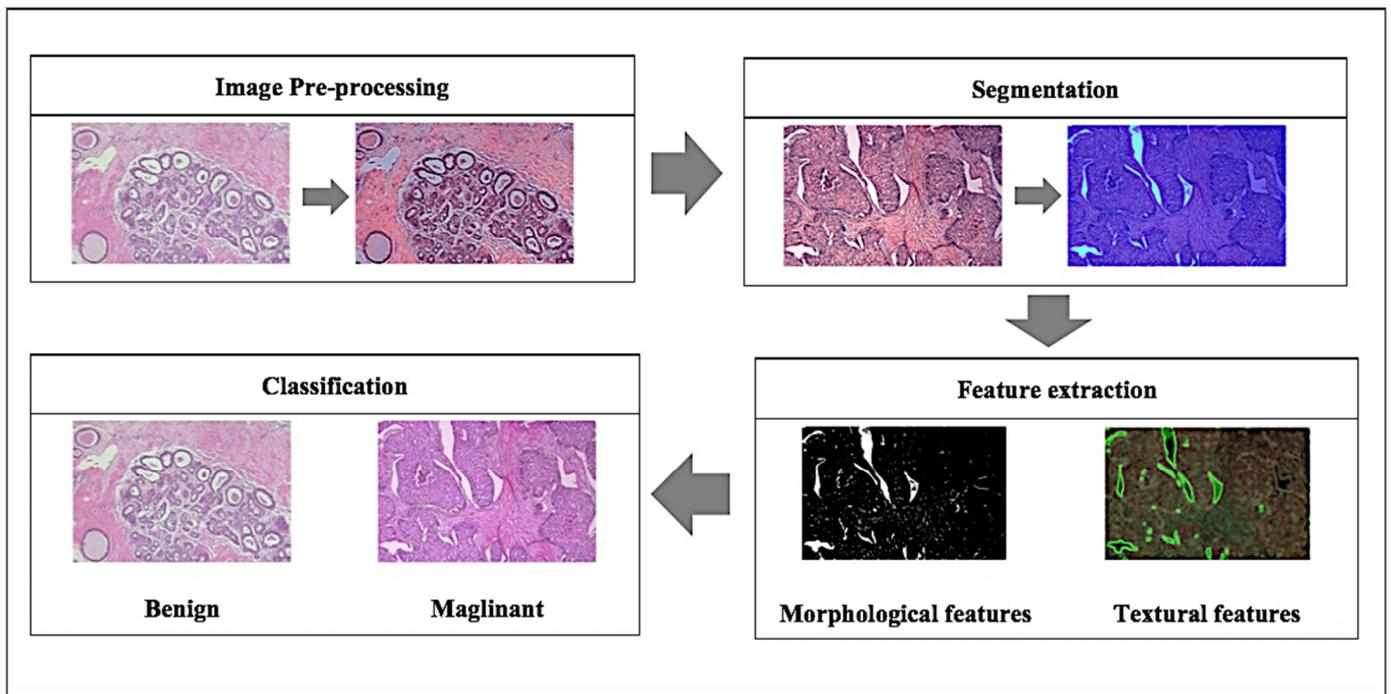


Figure 2. Core steps involved in a breast cancer computer-aided diagnosis system.

Breast cancer varies based on which part of the breast tissue becomes cancerous. Commonly, breast cancer starts in the cells that line the ducts of the breast; however, it may also grow in different areas of the breast such as the lobules, milk ducts or sometimes in between tissues, as illustrated in Figure 3 [9].

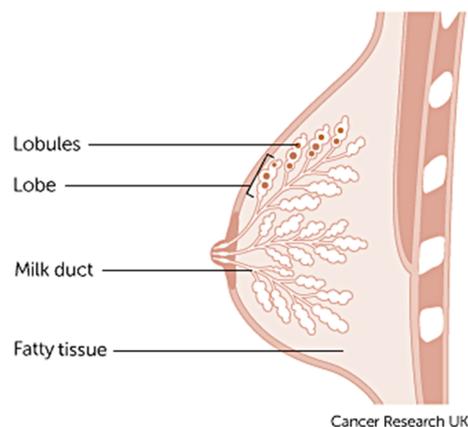


Figure 3. Anatomy of the breast credits to Cancer Research UK [9].

The term ‘breast cancer’ refers to a malignant tumour that has developed from cells in the breast that are considered cancerous and cause danger to health. The stage of this cancer is usually expressed as a number on a scale of 0 through IV, with stage 0 describing non-invasive cancers that are still within their original location and stage IV describing invasive cancers that have spread outside the breast [10]. In cases where cancer is detected, but no cancer cells are visible in the lymph glands, the breast cancer is of a lower risk. When spreading occurs, it carries a substantial risk of death, meaning that the cancer cells from the breast tissue have broken away, which can be carried to nearby lymph nodes by the lymph fluid (fluid that gathers waste products and drains into veins to be removed) [10]. Figure 4 demonstrates the lymph nodes around the breast [9].

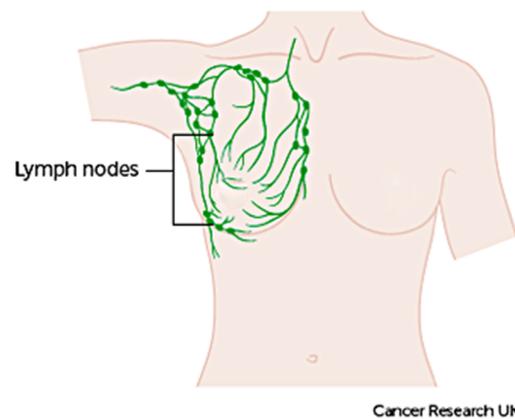


Figure 4. Network of lymph nodes around the breast based on a graphic created by Cancer Research UK [9].

Breast cancer can be distinguished as benign (non-cancerous) and malignant (cancerous/metastatic) tumours. Benign tissue refers to changes in normal tissue of breast parenchyma, which does not relate to the development of malignancy [11]. Contrarily, malignant tissue can be categorised into two types: in-situ carcinoma and invasive carcinoma. Additionally, in some cases benign breast tumours can be further divided into four subclass types, adenosis, fibroadenoma, phyllodes tumour, and tubular adenoma, whereas malignant breast tumours can be further divided into ductal carcinoma, lobular carcinoma, medullary carcinoma, mucinous carcinoma, tubular carcinoma, and papillary carcinoma [12].

Histopathology (histology) image samples of breast lesions are obtained through either needles or surgical operation, which are then later processed and allocated to a glass slide to undergo a staining process. Haematoxylin and eosin (H&E) and immunohistochemistry (IHC) are the most used histopathology staining protocols [4]. This development of scanners have digitalized histopathological tissue sections and turned digital pathology into a routine practice [13]. Currently, histopathological images play a vital role in cancer diagnosis because of the large amount of information they provide for medical image analysis [14]. Whole-slide images (WSI) can have multiple regions of breast lesion tissue, whereas microscopy images are patches derived from WSI, each representing one type of breast lesion only. In this paper we have chosen to study histopathology images of breast cancer in developing a machine learning based CAD system. Figure 5 demonstrates eight classes of breast cancer from the BreaKHis dataset [15].

The main contribution of this paper is to discuss the process, methods, comparisons, and remarks on developing a CAD expert system for breast cancer. The rest of the research paper is organized as follows: Section 2 explains the publicly available datasets for breast cancer histopathology images. The process of using a computer-aided expert system using histopathology images is presented in Section 3, which includes techniques employed in (1) image pre-processing, (2a) conventional CAD methods that employ segmentation, feature extraction, feature selection (dimension reduction) and classification; (2b) deep-learning-based CAD and (3) Performance evaluation. Finally, Section 4 discusses the conclusion and future directions for researchers are given Section 5.

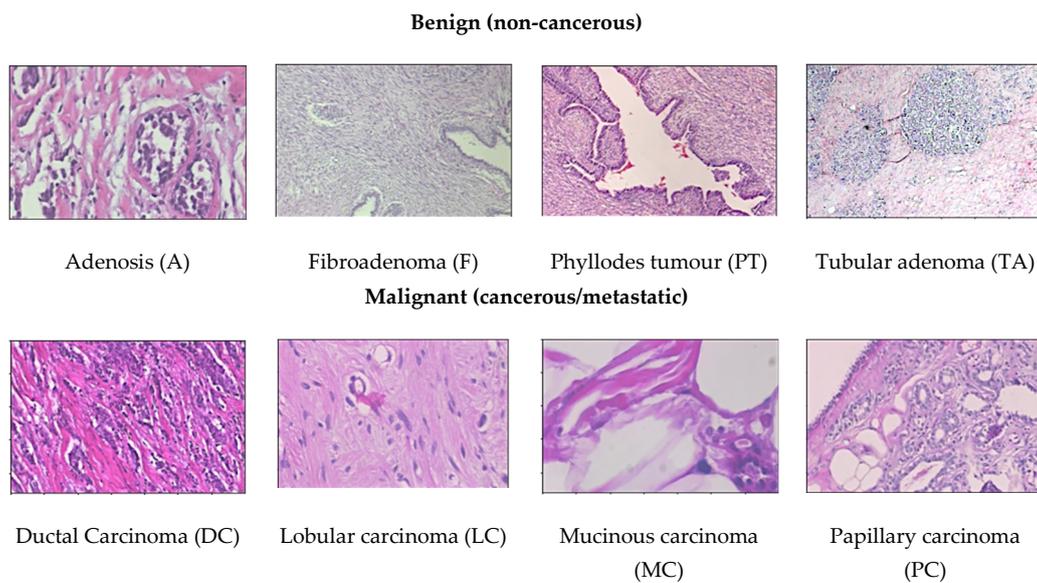


Figure 5. Sample of distinct types of breast cancer histopathology images from the BreakHis dataset [15].

2. Datasets for Breast Cancer Classification

In the field of medical image analysis, machine learning methodologies applied for histopathological images are developing rapidly. However, there is still a demand for an automatic system to get efficient and highly accurate results [14]. To obtain a large and representative annotated dataset to develop a machine learning method for CAD system is a challenging task [16]. Recently, there has been a rise in public challenges for breast cancer diagnosis which has attracted many researchers to this area of study. This section describes various publicly accessible datasets to assist future research and development.

1. BreakHis dataset [15]: This dataset provides 4 different magnification levels of $40\times$, $100\times$, $200\times$, and $400\times$ histology images of size 752×582 pixels. It consists of a total number of 7909 images acquired from a clinical study from January 2014 to December 2014 in P&D Laboratory, Brazil by 82 patients. For binary classification, there are two categories of benign and malignant to determine cancer or non-cancerous. There are 1995 images (652 benign and 1370 malignant) in $40\times$ magnification level; 2081 images (644 benign and 1437 malignant) in $100\times$ magnification level; 2013 images (623 benign and 1390 malignant) in $200\times$ magnification level; and 1820 images (588 benign and 1232 malignant) in $400\times$ magnification level in the dataset. To further perform multiclassification, the dataset contains four distinct types for each breast tumours. The category benign type of breast tumour consists of adenosis (A), fibroadenoma (F), phyllodes tumour (PT), and tubular adenoma (TA). The malignant type of breast tumour consists of ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary carcinoma (PC). This dataset is the most used dataset by many researchers for CAD breast cancer in histopathology images [11,17–30]. This dataset can be obtained from <https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/> (accessed on 16 March 2021).
2. Bioimaging Challenge 2015 dataset [31]: This dataset contains 269 images of haematoxylin and eosin (H&E)-stained breast cancer histology images with image size of 2048×1536 pixels. Images are provided in $200\times$ magnification level. For binary classification, there are two categories to determine cancer or non-cancerous. To further classify, the non-cancerous categories can be categorized as normal and benign, while the cancerous ones can be categorized as in situ carcinoma and invasive carcinoma. The training set has a total of 249 images form by 55 normal class, 69 benign class, 63 in situ carcinoma class, and 62 invasive carcinoma class, while the test set has a total of 20 images with 5 images for each class. Additionally, there is an extended

- test set with more diversity provided with a total of 16 images available. In this extended test set, there are 4 images for each class. This dataset can be obtained from <https://rdm.inesctec.pt/dataset/nis-2017-003> (accessed on 16 March 2021).
3. BACH (BreAst Cancer Histology) dataset [32]: The ICIAR 2018 challenge resulted in the BreAst Cancer Histology (BACH) image dataset, which is an extended version of the Bioimaging 2015 breast histology classification challenge dataset with similar image sizes and magnification levels [31]. The dataset has a total number of 400 images, respectively classified to a total number of 100 normal class, 100 benign class, 100 in situ carcinoma class, and 100 invasive carcinoma class. The test set has a total of 100 images without any labels. The dataset can be obtained from <https://iciar2018-challenge.grand-challenge.org/> (accessed on 16 March 2021).
 4. CAMELYON dataset [33]: The Cancer Metastases in Lymph Nodes Challenge breast cancer metastasis detection dataset combines two datasets collected from CAMELYON16 and CAMELYON17 challenges, with each image approximately 1×105 by 2×105 pixels at the highest resolution. The first dataset CAMELYON16 consists of a total 400 whole-slide images (WSIs) of haematoxylin and eosin (H&E)-stained lymph node sections collected from Radboud University Medical Center (Nijmegen, The Netherlands) and the University Medical Center Utrecht (Utrecht, The Netherlands). Each image is annotated with a binary label for classification, showing normal and presence of metastases tissue. There are two sets of training datasets, the first has a total number of 170 images, formed of 100 normal class and 70 metastases class, while the second has a total number of 100 images formed of 60 normal class and 40 metastases class. The test set holds a total number of 130 images. The CAMELYON17 dataset consists of a total of 1399 histology breast images. This version is extended from the CAMELYON16 which include patients testing for breast cancer from the CAMELYON16 challenge with an additional three medical centres from the Netherlands, specifically: slides from 130 lymph node resections from Radboud University Medical Center in Nijmegen (RUMC), 144 from Canisius-Wilhelmina Hospital in Nijmegen (CWZ), 129 from University Medical Center Utrecht (UMCU), 168 from Rijnstate Hospital in Arnhem (RST), and 140 from the Laboratory of Pathology East-Netherlands in Hengelo (LPON) [34]. The dataset can be obtained from <https://camelyon17.grand-challenge.org> (accessed on 16 March 2021).
 5. PatchCamelyon (PCam) dataset [35]: Whole slide images (WSI) are computationally expensive and only require the small regions of interest (ROIs) from the entire image, therefore it would need to estimate a significantly substantial number of parameters. Thus, this version of dataset is derived from the CAMELYON dataset with a total number of 327.680 histopathologic scans of lymph node sections images, each in the size of 96×96 px pixels. Like the CAMELYON dataset, each image is annotated with binary label for classification, showing normal and presence of metastases tissue. The main difference and advantage of this dataset is that it is bigger than CIFAR10, smaller than ImageNet, additionally it is trainable on a single GPU to able to achieve competitive scores in the CAMELYON16 tasks of cancer detection and WSI diagnosis. PCam contributed by supplying the segmented tissue parts that separated tissue and background from the whole slide images. The dataset can be obtained from <https://github.com/basveeling/pcam> (accessed on 16 March 2021).
 6. MITOS-12 dataset [36]: The conference ICPR 2012 supplied the MITOS dataset benchmark that consists of 50 histopathology images of haematoxylin and eosin (H&E)-stained slides of breast cancer images from 5 different breast biopsies at $40\times$ magnification level. However, this dataset is too small to produce an exceptionally reliable performance and the robustness of the diagnosis system is limited. Therefore, an extended version of the dataset (MITOS-ATYPIA-14) was presented at ICPR 2014.
 7. MITOS-ATYPIA-14 dataset [37]: The grand challenge dataset was presented at the ICPR 2014 conference, extended from the MITOS-12 challenge that provides haematoxylin and eosin (H&E)-stained slides of breast cancer images with the size of

1539 × 1376 pixels at 20× and 40× magnification level. There is a training set with a total number of 1200 images acquired from 16 different biopsies and testing set with a total number of 496 images acquired from 5 different breast biopsies. The dataset consists of a significantly diverse variation of stained images in many conditions to elevate the challenge to achieve a more exceptional performance. The dataset can be obtained from <https://mitos-atypia-14.grand-challenge.org/> (accessed on 16 March 2021).

8. TUPAC16 dataset [38]: The dataset consists of a total number of 73 breast cancer histopathology images at 40× magnification level from three pathology centres in the Netherlands. The dataset is composed of 23 test images with a size of 2000 × 2000 pixels and 50 training images with a size of 5657 × 5657 pixels collected from two separate pathology centres. The images contained in the training dataset are later cropped randomly to the size of 2000 × 2000 pixels. The dataset can be obtained from <http://tupac.tue-image.nl/node/3> (accessed on 16 March 2021).
9. UCSB bio segmentation benchmark (UCSB-BB) [39]: This dataset contains 50 haematoxylin and eosin (H&E)-stained histopathology images used in breast cancer cell detection with the size of 896 × 768 pixels and a ground truth table. Each image is annotated with binary label for classification, it contains half malignant class and half benign class. The dataset can be obtained from <https://bioimage.ucsb.edu/research/bio-segmentation> (accessed on 16 March 2021).

3. Computer-Aided Diagnosis Expert Systems

CAD systems have not only produced faster diagnosis results but have also emerged as an additional opinion to assist pathologists to avoid overlooking abnormal features. This automated solution can be explained in two sub-categories:

1. Computer-aided detection (CADe) systems, which detect cancer or metastatic tissue.
2. Computer-aided diagnosis (CADx) systems, which determine the distinct types of breast cancer.

There are two approaches in developing a CAD system which are the conventional method and deep learning method. The main difference between these two types of methods is that conventional CAD methods are a traditional approach of extracting the features from an image based on human-defined descriptors to perform classification. Deep learning CAD methods are types of automated learning that can discover representations of data automatically by transforming the input information into multiple layers of abstractions [8]. Figure 6 illustrates these two methods for CAD systems.

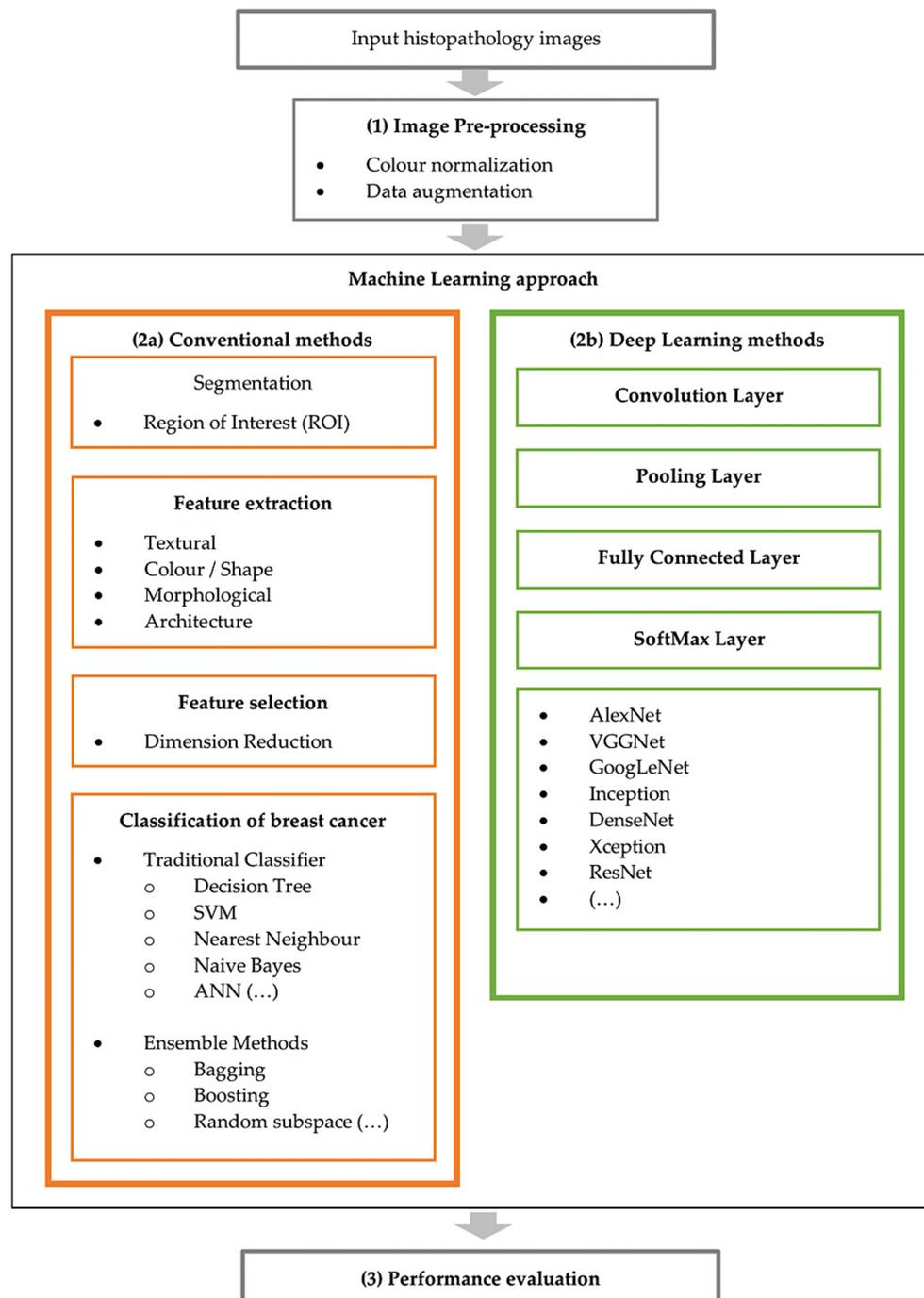


Figure 6. Overview structure of methods used to classify breast cancer in CAD systems.

3.1. Image Pre-Processing

Image pre-processing is an effective route to apply as data preparation at the first step to make raw data more suitable for further analysis. In the case of histopathology images, the most used pre-processing technique is colour normalisation because of the colour variation obtained in these types of images and the powerful impact on the machine learning model. Data augmentation is another commonly used technique for a small dataset. In this section, the techniques of (1) colour normalisation and (2) data augmentation are presented.

- (1) Colour normalisation: The inconsistent various appearances of stained sections is amongst the foremost challenges to analyse histopathological images [40]. This is

because the samples are collected under various inconsistent conditions of tissue slices, preparation or image acquisition, noise arising, lightning conditions, and protocols of staining while capturing the digital image [40]. Therefore, these variations could produce samples with different colour intensities [41]. Research studies [18,42] have shown the significant effect of stain normalisation that enhances the performance of breast cancer classification. Here, a few colour normalisation techniques will be investigated by categorizing them into three types of method which are global colour normalisation, the supervised method, and the unsupervised method for stain separation.

- Global colour normalisation: This method is suitable for histology images due to comprehensible values of autocorrelation coefficient or spatial dependency of pixel (intensity). This method separates colour and intensity information using principal component analysis (PCA) [43]. Reinhard et al.'s method was one of the first techniques, which uses a simple statistical analysis to achieve colour correction by comparing one image's colour boundaries and choosing it as an appropriate source image as a benchmark, applying it as characteristic to all the other images [43]. It uses an unsupervised method to heuristically estimate the absorbance coefficients for the stains for every image and the staining concentrations for every pixel to recompose the images [43].
- Supervised method for stain separation: In this method, images are converted to optical density (OD) space due to Beer's law [44] that suggests colour stains act linearly in OD space, given in Equation (1).

$$V = \log\left(\frac{I_0}{I}\right) \quad (1)$$

where V represents the intensity in OD space, I represents the intensity in RGB space, and I_0 represents the illuminating intensity incident on the sample [45]. Khan et al. proposed a method to use stain colour descriptors to compute image-specific stain matrices for stain normalisation [46]. Then, stain separation is applied to obtain different stain concentration values from the image and provide a nonlinear (spline based) mapping function; meanwhile all images will be replaced using the normalised stain channels [46].

- Unsupervised method for stain separation: Training is not required because it is expected to learn itself [47]. Macenko et al. first proposed a method to use singular value decomposition method (SVD) to obtain optical density of images to perform quantitative analysis-based colour normalisation [48]. Kothari et al. then proposed a method based on histogram specification using the quantile normalisation based on distinct colour channels obtain from images to match each image to the target image histogram colour channels [49]. Bejnordi et al. later proposed an improved version which relies solely on colour features; their algorithm makes use of spatial information to achieve robustness against severe colour and intensity variations [50]. The comparison of colour normalisation methods is provided in Table 1.

Table 1. Comparison of colour normalisation methods.

Ref	Proposed Approach	Method	Advantages	Disadvantages
[43]	Colour transfer algorithm	<ol style="list-style-type: none"> 1. Convert the colour space of an image from RGB to $\alpha\beta$ [51]. 2. Transform the background colour of images based on the target colour space. 3. Convert images back to RGB colour space. 	<ul style="list-style-type: none"> • All images will have the same consistent range of contrast. • Structure of images remains. 	<ul style="list-style-type: none"> • Stains in images are not separated properly due to the type of colour space conversion ($\alpha\beta$).
[48]	Fringe search algorithm	<ol style="list-style-type: none"> 1. Convert the colour space of an image from RGB to $\alpha\beta$ [51]. 2. Create plane based on calculated two largest SVD. 3. Estimate data onto that plane. 4. Search for corresponding DOF angles. 5. Robust predictions of minimum and maximum are calculated by the αth and $(100-\alpha)$th percentile. 6. Convert these obtain DOF angles values back to OD space. 	<ul style="list-style-type: none"> • Negative coefficient is not found in colour appearance matrix. • Absence of ambiguity. 	<ul style="list-style-type: none"> • Not ideal for automated tumour detection algorithm because the DOF angle values are estimated observationally. • Original images are not preserved.
[49]	Automated colour segmentation algorithm	<ol style="list-style-type: none"> 1. Apply pre segmentation by extracting the unique colours in the image to obtain colour map. 2. Include knowledges from pre segmented reference images to normalise. 3. Apply voting scheme to evaluate on preliminary segmentation labels. 4. Apply segmentation to new images with the multiple reference images and combine labels from previous step. 	<ul style="list-style-type: none"> • High accuracy. • Robustness. • Makes use of expert domain knowledge. • Retains the morphology of images. 	<ul style="list-style-type: none"> • Colour map histogram distortion due to chromatic aberration. • Restricted to segmentation problems with four stain colours.
[46]	Nonlinear mapping approach	<ol style="list-style-type: none"> 1. Map both target image and source images to a representation, where each channel relates to a separate chemical stain. 2. Calculate the statistics of each corresponding channel by learning a supervised classification method (RVM). 3. Apply a nonlinear correction (mapping) to normalise each separate channel based on previous calculation. 4. Reconstruct the normalised source image using the normalised stain channels. 	<ul style="list-style-type: none"> • Satisfactory performance overall for separating stains. • Performs at pixel level to achieve superior performance. 	<ul style="list-style-type: none"> • High computation complexity. • Using nonlinear correction (mapping) functions might destroy the original image structure i.e., colour histogram. • Impossible to convert back to original form of an image after mapping.

Table 1. Cont.

Ref	Proposed Approach	Method	Advantages	Disadvantages
[50]	Whole-slide image colour standardiser (WSICS) algorithm	<ol style="list-style-type: none"> 1. Apply hue-saturation-density (HSD) colour transformation [52] to obtain two chromatic components and a density component 2. Gather distribution of the transformation of haematoxylin and eosin (H&E). 3. Calculate weight contribution of stain in every pixel. 4. Convert HSD back to RGB. 	<ul style="list-style-type: none"> • Robustness. • Remain spatial information of images. • It is an unsupervised method capable of detecting all stain components correctly. 	<ul style="list-style-type: none"> • Losing the original background colour during the process. • High processing time. • Not all information of images is preserved.

Recently, due to the significant performance and stability portrayed by these proposed methodologies for colour normalisation, many researchers have adapted these popular proposed methods above as part of their colour normalisation processes. Table 2 demonstrates the methodology used by several recent research studies for breast cancer CAD systems.

Table 2. Comparison of adopted colour normalisation methods.

Ref	Year	Adapted Colour Normalisation Method
[42]	2017	Bejnordi et al. and Macenko et al. methodology
[31]	2017	Macenko et al. methodology
[27]	2018	Reinhard et al. and Kothari et al. methodology
[53]	2018	Macenko et al. methodology
[18]	2019	Macenko et al. methodology
[19]	2019	Macenko et al. methodology
[54]	2019	Reinhard et al. methodology
[29]	2019	Macenko et al. methodology
[55]	2019	Simplified version of Bejnordi et al. methodology
[22]	2021	Khan et al. methodology

- (2) Data augmentation: A data-space solution to the problem of limited data by enhancing the size of training datasets to generate a better learning model [56]. Tellex et al. showed that to obtain a particularly reliable performance of CAD system on histopathology images, colour normalisation should be used along with data augmentation [57]. This procedure will imply data wrapping and oversampling over the dataset to increase the sample size of the training dataset as a limited dataset and overfitting is a common challenge [56]. These processed include various image transformations to modify the image morphology [57,58]. If we were to look at one image from a single perspective and make a determination, it is more likely to be prone to error compared to if we were to look at it from several perspectives to make the final determination. Taking this into breast cancer analysis, checking the image with several more perspectives provides a more confident and accurate answer to which class it belongs to. Thus, this procedure provides a broader interpretation to the original image. The comparison of data augmentation techniques applied by several research studies is provided in Table 3.

Table 3. Comparison of data augmentation.

Ref	Flipping	Cropping/Shearing	Rotation	Translation	Shifting	Scaling	Zooming	Contrast	Fill Mode	Brightness
[11]	✓	✓	✓	✓	✓		✓			
[27]	✓		✓		✓					
[18]	✓	✓	✓	✓						
[53]		✓				✓				
[19]	✓	✓	✓				✓	✓	✓	
[17]	✓	✓	✓							
[21]	✓		✓		✓					✓
[26]	✓		✓							
[59]	✓	✓		✓		✓				
[60]			✓							
[20]	✓	✓	✓		✓				✓	

3.2. Conventional CAD Methods

3.2.1. Segmentation

The segmentation process takes part to locate the edges and boundaries of regions in a histopathology image to extract the cells in the images. It can be crucial to identify the region of interest (ROI) and highlight these significant regions in the images. This procedure involves partitioning the image I into non-overlapping regions [61,62], as seen in Equation (2).

$$\cup I_i = I \quad \text{and} \quad I_i \cap I_j = \emptyset, \quad \text{where } I \neq j \tag{2}$$

Table 4 provides a summary for each commonly used segmentation technique along with definition, advantages, and limitations. Figure 7 illustrates a general overview of the approached techniques with some examples.

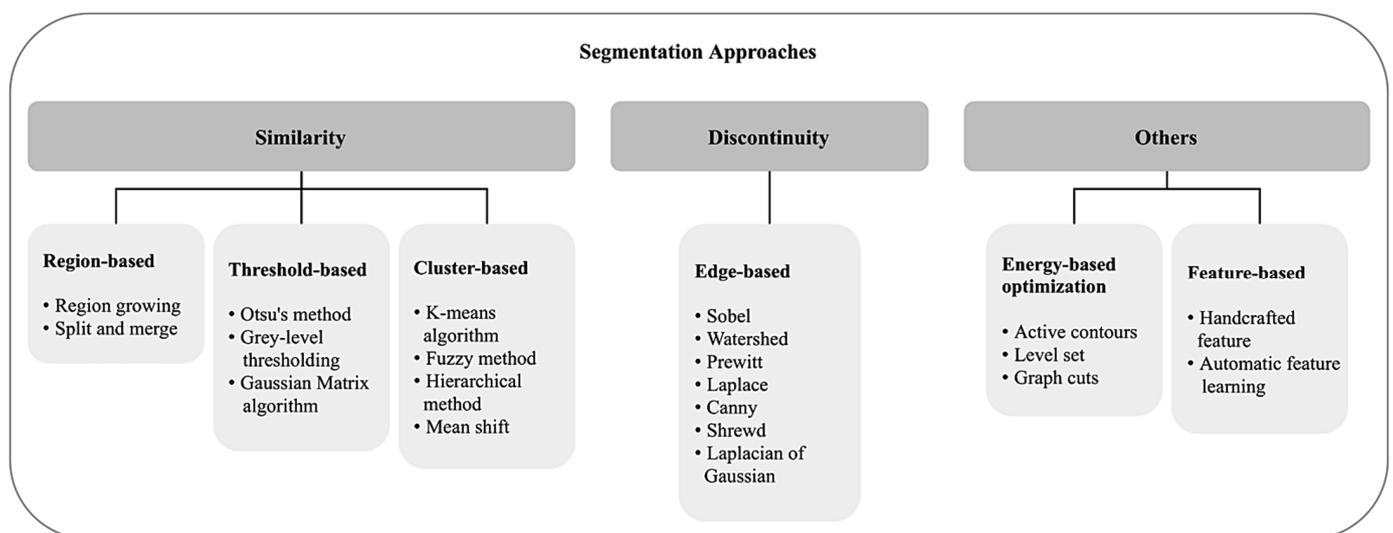


Figure 7. General view of segmentation approaches and techniques.

Table 4. Summary of segmentation approaches along with their definition, advantages, and limitations.

Approach	Segmentation Technique	Definition	Advantages	Limitations
Image processing	Region-based	Each pixel will be separated into groups (regions) in a homogeneous way based on a seed point.	<ul style="list-style-type: none"> Parameters are easy to be adjusted. Enables segmentation for multiple class. Beneficial for noisy images as their edges will be harder to detect. 	<ul style="list-style-type: none"> Local region solution. Computationally expensive. Not robust enough. Missing/weak boundaries. Need to specify seed point. Output varies with different seed point.
	Edge-based	Edges are defined based on the sharp discontinuity (i.e., intensity) in the image.	<ul style="list-style-type: none"> Low computational complexity. Simple technique. Works fine with images with prominent edges. 	<ul style="list-style-type: none"> Prone to over-segmentation error. Requires further morphological operation tuning for accurate result. Cannot apply on images with smooth edges. Requires high quality images. Hard to interpret with noisy images.
	Thresholding-based	Transform every pixel based on a threshold value obtained from a histogram of image that corresponds to regions.	<ul style="list-style-type: none"> Stable and flexible. Easy to implement. 	<ul style="list-style-type: none"> Dependant on selection of an effective/correct threshold value. Not suitable for histology images because of the high complexity and various intensity distributions in images.
Machine learning	Cluster-based	Objects in image will be categorised into specific regions (groups) based on their similarity in pixels.	<ul style="list-style-type: none"> Efficient. Easy to implement. Solution dependant on the initialization. Able to preserve the information. Suitable for microscopic biopsy images. 	<ul style="list-style-type: none"> Need to specify number of clusters. Sensitive to outliers.
	Energy-based optimization	Contour object of interest by minimizing/maximizing a predefined cos function.	<ul style="list-style-type: none"> High accuracy. Robustness. 	<ul style="list-style-type: none"> High complexity. High computational time. Requires defining an effective cost function.
	Feature-based	Uses a model to train and learn the features to determine which pixels are ROI.	<ul style="list-style-type: none"> Supervised learning method. Robustness. 	<ul style="list-style-type: none"> Application dependant.

- Region-based segmentation: There are two main techniques which are (1) region growing and (2) region splitting and merging. Rouhi et al. proposed the application of an automated region growing for segmentation on breast tumour histology images by using an artificial neural network (ANN) to obtain a threshold [62]. Rundo et al. used split and merging algorithms based on the seed selection by an adaptive region growing procedure [63]. Lu et al. applied a multi-scale Laplacian of Gaussian (LoG) [64] to detect the seed points and feed the filtered image to a mean-shift algorithm for segmentation, followed by some morphological operations [65].
- Edge-based segmentation: To obtain critical properties, this structural technique can be implemented in several methods for recognising the edges, such as Sobel [66], Watershed [67], Prewitt [68], Laplace [69], Canny [70], and LoG [64]. This process is illustrated in Figure 8.



Figure 8. Overall process of edge-based segmentation.

George et al. applied the watershed method to extract shape and texture features of nuclei, where both features contribute to the training of an accurate nuclei classifier for breast cancer [71]. Faridi et al. used the Distance Regularized Level Set Evolution (DRLSE) algorithm for segmentation; the process includes morphological operations to detect centre of nuclei and Difference of Gaussian (DoG) filtering [72] was applied to extract nuclear boundaries [73].

- Threshold-based segmentation: To produce a less complex image, the main concept is to transform every pixel based on a threshold value; any pixels with intensity less than a threshold value/limit esteem T (constant value) will be replaced with black pixels (0), otherwise replaced with white pixels (1). The input image $g(x, y)$ transformed to a threshold image $f(x, y)$ can be represented mathematically as shown in Equation (3).

$$f(x, y) = \begin{cases} 1, & \text{if } g(x, y) > T \\ 0, & \text{if } g(x, y) \leq T \end{cases}, \text{ where } T \text{ is a threshold value} \quad (3)$$

A few popular techniques applied in this approach are Otsu thresholding [74], grey-level thresholding [72], and gaussian matrix thresholding [72]. Zarella et al. proposed a scheme to segment breast nuclei from other parts of the cell using Otsu thresholding [75]. Saha et al. proposed an automatic nucleus segmentation on the image using histogram-based thresholding with a result of 97% accuracy in nucleus detection [76]. Moncayo et al. used Maximally Stable Extreme Regions (MSER) to perform segmentation on nuclei regions on the image's haematoxylin contribution map, in which several thresholds are applied to the image and areas that change minimally are identified as MSER, followed by some further morphological operations [77]. A novel approach was proposed by Khairuzzaman and Chaudhury to apply a multilevel thresholding based on Grey Wolf Optimizer (GWO) using Kapur's entropy and Otsu's between class variances functions [78]. Sirinukunwattana et al. proposed a thresholding method to group intensity features represented by a sparse coding to create a dictionary [79].

- Cluster-based segmentation: This can be described in two clustering methods, hierarchal and partitioning [80]. Hierarchal clustering performs recursively to explore nested clusters in agglomerative (bottom to up) or divisive (top to down) ways [80], whereas partitioning clustering iteratively divides into hard clustering and fuzzy clustering [81]. Kowal et al. applied a cluster approach algorithm for nuclei segmentation

from biopsy microscopic images, and achieved a high classification accuracy [82]. Kumar et al. used a k-means clustering based segmentation algorithm and mentioned that this method performs better in comparison to other commonly used segmentation methods [83]. A two-step k-means was applied for segmentation by Shi et al. to consider local correlation of pixels; they first generate a poorly segmented cytoplasm, then in a second step the segmentation does not take into account the nuclei identified during the first clustering; finally, a watershed transform was applied to complete the segmentation [84]. Maqlin et al. suggested a segmentation method based on k-means clustering algorithm to recover the missing edge boundaries based on a convex grouping algorithm, which was suitable for open vesicular and patchy types of nuclei that are commonly obtained in high-risk breast cancers [85].

- Energy-based optimization: This technique defines a cost function, and the process will minimize/maximize the function based on the object of interest (ROI) in the images. A study by Belsare et al. used a spatio-colour-texture graph cut segmentation algorithm to perform segmentation as epithelial lining surrounding the lumen [86]. Wan et al. used a combination of boundary and region information to perform a hybrid active contour method to achieve an automated segmentation of the nuclear region [87], where the energy function was defined as Equation (4).

$$\varepsilon(\phi) = -\alpha \int_w (Z - \mu)H(\phi)d\omega + \beta \int_w G|\nabla H(\phi)|d\omega \quad (4)$$

where Z is the image to be segmented, $H(\phi)$ denotes the Heaviside function, ω represents the image domain, $G = G(|\nabla Z|)$ is the gradient of the image, and α and β are pre-defined weights for the balancing of the two terms. Zhang et al. proposed a three-phase level set method to set contour segments into groups and achieved high accuracy [88]. Jia et al. used a rough segmentation method to combine watershed and improved Gradient Vector Flow (GVF) Snake model to separate nuclei/cells in an image from the background to enhance the segmentation accuracy [89].

- Feature-based segmentation: Automatic segmentation based on feature learning has been commonly used for analysing medical images [90]. Song et al. used a multi-scale convolutional network to accurately apply segmentation of cervical cytoplasm and nuclei [91]. Xu and Huang applied a distributed deep neural network architecture to detect cells [92]. Rouhi et al. also proposed a cellular neural network (CNN) to perform segmentation by using genetic algorithm (GA) to determine the parameters [62]. Graham et al. proposed a deep learning method called the HoVer-Net which is a network that targets simultaneous segmentation and classification of nuclei based on the horizontal and vertical distance maps to separate clustered nuclei [93]. Zarella et al. trained an SVM model to learn the features to distinguish between stained pixels and unstained pixels using HSV colour space to identify regions of interest [94]. A summary of different segmentation approaches by several researchers is provided in Table 5.

Table 5. Summary of different segmentation approaches by several researchers.

Type of Technique(s) Employed	Ref	Year	Approach(es)	Remarks
Cluster-based	[82]	2013	Cluster algorithm	<ul style="list-style-type: none"> Contributed to a good feature extraction result.
	[83]	2015	K-means clustering algorithm	<ul style="list-style-type: none"> Able to preserve the desired information. Best suited for microscopic biopsy images.
	[84]	2017	Two step k-means clustering, and watershed transform	<ul style="list-style-type: none"> Performed on lymph nodes histology images. Considers local correlation of every pixel.
	[85]	2015	Segmentation: k-means clustering algorithm. Recover edges: Convex grouping algorithm.	<ul style="list-style-type: none"> Produces incorrect clusters when an image has fewer pixels of nuclei. Does not achieve splitting of overlapped cells.
Edge-based	[71]	2014	Watershed	<ul style="list-style-type: none"> Less complex and more computationally efficient. Prone to over-segmentation.
	[73]	2016	Distance regularized level set evolution (DRLSE) algorithm	<ul style="list-style-type: none"> Not sensitive to parameters. Able to detect and segment overlapping cells.
Energy-based optimization	[86]	2016	Graph cut: Spatio-colour-texture graph segmentation algorithm [95].	<ul style="list-style-type: none"> A novel integrated method. Similarity based super pixel generation method.
	[87]	2017	Hybrid active contour method	<ul style="list-style-type: none"> Accurate segmentation. Not sensitive to parameters.
	[88]	2017	Three-phase level set method to set contour	<ul style="list-style-type: none"> High accuracy in both clear and blurry nuclei images.
	[89]	2021	Watershed and improved gradient vector flow (GVF) snake model	<ul style="list-style-type: none"> Powerful segmentation model. Less prone to overlapping or obstructed boundaries

Table 5. Cont.

Type of Technique(s) Employed	Ref	Year	Approach(es)	Remarks
Feature-based	[94]	2015	Support vector machine (SVM)	<ul style="list-style-type: none"> Manually selects positively and negatively stained pixels from a set of representative images.
	[91]	2015	Multi-scale convolutional network	<ul style="list-style-type: none"> Fully automated segmentation process.
	[92]	2016	Distributed deep neural network	<ul style="list-style-type: none"> Fully automated segmentation process. High sensitivity of preserved images.
	[62]	2016	Cellular neural network (CNN) trained on genetic algorithm (GA) parameters	<ul style="list-style-type: none"> Prior information used to reduce errors.
	[93]	2019	Deep learning using HoVer-Net	<ul style="list-style-type: none"> Based on horizontal and vertical distance maps.
	[59]	2020	Faster R-CNN	<ul style="list-style-type: none"> A new method that has not been fully explored in breast cancer segmentation of mitosis cell. Computationally inexpensive. High accuracy.
Region-based	[62]	2015	Automated region growing using ANN to obtain threshold	<ul style="list-style-type: none"> Efficiently select threshold value to reduce errors.
	[65]	2015	Mean-shift algorithm	<ul style="list-style-type: none"> Application dependant. Able to handle arbitrary feature spaces.
	[63]	2016	Split and merging algorithm based on adaptive region growing	<ul style="list-style-type: none"> Requires a decent quality image. Fast and accurate when image is in good condition. Able to handle noise in image.

Table 5. Cont.

Type of Technique(s) Employed	Ref	Year	Approach(es)	Remarks
Threshold-based	[77]	2015	Maximally Stable Extreme Regions (MSER)	<ul style="list-style-type: none"> • Accurate segmentation on complicated images.
	[78]	2017	Multilevel thresholding based on Grey Wolf Optimizer (GWO) algorithm using Kapur's entropy and Otsu's between class variance functions.	<ul style="list-style-type: none"> • More stable and yields solutions. • Performs faster than BFO, however slower than the PSO-based method.
	[76]	2015	Histogram-based thresholding	<ul style="list-style-type: none"> • Hard to determine a suitable threshold.
	[79]	2015	Dictionary, thresholding	<ul style="list-style-type: none"> • This is a mitotic cell detection system using a dictionary of cells.
	[75]	2017	Otsu thresholding	<ul style="list-style-type: none"> • Does not require definition of many parameters.

3.2.2. Feature Extraction

Feature extraction is one of the essential steps to pick out a set of features that contain the most effective, relevant, and discriminating information and characteristics of ROI/entire images to be employed for classification. Overall, we can divide the image feature descriptors into three dimensions (shape, pattern and spectra, and density). From Figure 9, we can observe a feature taxonomy based on feature descriptor dimensions from the 3D axis [96].

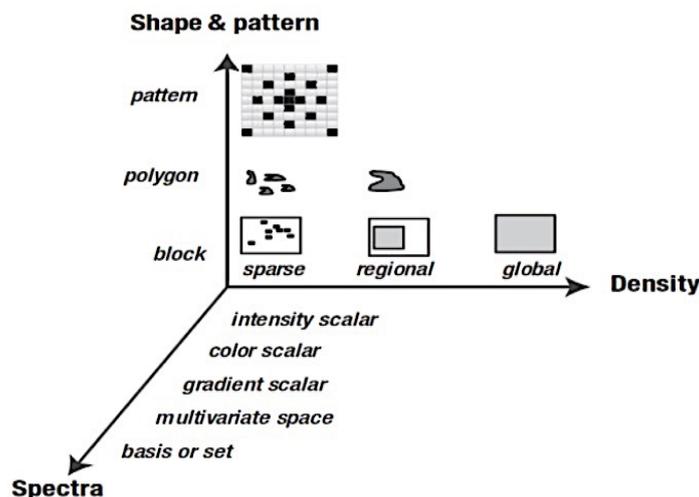


Figure 9. Taxonomy for feature descriptor dimensions created by Krig, S. [96].

In this section, feature extraction methods observed in the existing literature will be presented. Then, we will discuss the type of features extracted from images.

- **Morphological Features:** Describes the details of the image regarding information in geometric aspects such as the size (radii, perimeter, and area) and shape (smoothness, compactness, symmetry, roundness, and concavity) of a cell [97].
- **Textural Features:** Collects information of various intensity of every pixel value from histology images by applying several methods to obtain a number of properties such as smoothness, coarseness, and regularity [97].
- **Graph-Based Topological Features (architectural features):** Describes the structure and spatial arrangement of nuclei in a tumour tissue [97]. When dealing with histopathological images, the arrangement and shape of nuclei is connected to the cancer development, therefore this architecture may be calculated using graph-based techniques [98,99]. There are many different topology-based features including the count of number of nodes, edges, edge length, and roundness factor to detect the tissues [100,101]. There are three types of common graph features: Voronoi diagram, Delaunay triangulation, and minimum spanning tree, as shown in Figure 10.

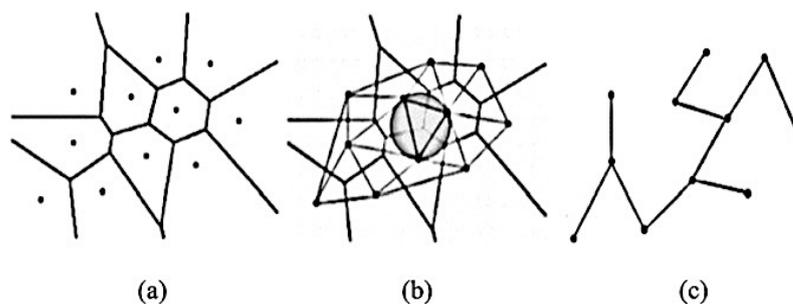


Figure 10. Types of graph-based topological features illustrated as (a) Voronoi diagram, (b) Delaunay triangulation, (c) minimum spanning tree [102].

Belsare et al. proposed to extract the textural features such as grey-level co-occurrence matrix (GLCM), graph run length matrix (GRLM) features, and Euler number; their system was able to achieve a 100% accuracy in 70 histopathological images on a dataset from Department of Pathology, Govt. Medical College and Hospital, Nagpur, India [86]. Balazsi et al. proposed an invasive ductal breast carcinoma detector that extracts patches and a set of 16,128 features derived from multiple histograms and LBP (multiple radii) using CIELAB, grey-scale and RGB colour spaces describes each patch on a dataset from McGill University Hospital Centre pathology registry [103]. Wan et al. extracted several lots of information on multi-level features set in regards of the pixel-, object-, and semantic-level features [87]. The pixel-based features are textural features, Kirsch filters, first-order features, Gabor filters, Haralick features, HoG, and LBP. The object-based features are architectural ones represented by graphs using Voronoi diagram (VD), minimum spanning tree (MST), and Delaunay triangulation (DT). Semantic-level features capture heterogeneity of cancer biology using convolutional neural networks (CNN)-derived descriptors on a dataset from China's No. 91 Central Hospital of PLA [87].

Recently, many authors have provided a wide range of publicly available breast cancer histopathological datasets to resolve the limitations, as shown in Section 2 Datasets for Breast Cancer Classification. Spanhol et al. provided the 'BreKHis' dataset and also performed some initial experiments by applying a handcrafted method to extract textural features like local binary patterns (LBP), completed LBP (CLBP), local phase quantization (LPQ), grey-level co-occurrence matrix (GLCM), threshold adjacency statistics (TAS), parameter-free threshold adjacency statistics (PFTAS), and one key point descriptor, named ORB [15]. Sudharshan et al. investigated the parameter-free threshold adjacency statistics (PFTAS) features feeding into a multiple-instance learning (MIL)-based nonparametric classifier; their results achieved the highest patient recognition rate (Prr) on the BreKHis dataset [30]. Gupta et al. used a set of colour-textural features including Gabor filters features, wavelet features, and local binary patterns (LBP) features to be fed into an ensemble classifier; their classification results achieved a 90.32% accuracy using the BreKHis dataset on $200\times$ magnification [104]. Chan and Tuszynski applied fractal dimension features for breast cancer detection; their results show that these features perform well at a magnification of $40\times$ to distinguish malignant and benign tumours on the BreKHis dataset [23].

Kumar et al. extracted various biologically interpretable and clinically significant shapes as well as morphology-based features, which include the grey-level texture features, colour-based features, colour grey-level texture features, Law's Texture Energy based features, Tamura's features, and wavelet features on a dataset with a total of 2828 histology images (histologyDS2828) [83]. Rezaeilouyeh et al. mentioned that wavelet features do not have directional sensitivity, which makes them unsuitable for detecting directional features, thus they proposed to use shearlets instead [105]. The authors proposed to perform calculation of shearlet coefficients to extract textural features to obtain them as primary features before feeding it to a CNN training stage for classification on the UCSB-BB dataset [105]. On the same dataset, Anuranjeeta et al. applied shape and morphological features of cells and achieved a result of 85.7% accuracy among 70 images [106]. On the same dataset, Bruno et al. applied a curvelet transform to handle multi-scale using a local binary pattern (LBP) algorithm to extract features from curvelet coefficients [107]. Moncayo et al. proposed a set of extracted features named bag of features (BoF) from the multi-scale textural features describing the segmented nuclei region, that is then assigned to the most similar atom of a previously learned dictionary using k-means algorithm on the National Cancer Institute: The Cancer Genome Atlas dataset [77].

Based on research, the most frequent applied hybrid techniques for cancer classification are combining morphological and textural features, for instance the work by [66,108]. Gandomkar et al. applied a hybrid approach of using segmentation-based and texture-based methods to extract features to obtain features that can discriminate between the different cancer classifications on the MITOS-ATYPIA-14 dataset [109]. Lu et al. extracted

a total combination of around 142 morphological and textural features, which included the size, mean, stain's standard deviation, sum, entropy, and mean of gradient magnitude image, 3 Tamura texture features, 44 grey-level run-length matrix-based textural features, and 88 cooccurrence matrix-based Haralick texture features on the MITOS-ATYPIA-14 dataset [65]. Khan et al. proposed to extract textural features of geodesic means of region covariance (RC) descriptors by calculating RC descriptors for different segmented regions, whereas a single descriptor for the whole image is derived by the geodesic geometric mean of these calculated RC on the MITOS-12 dataset [110]. Maroof et al. proposed a method of using hybrid feature space to combine colour features with morphological and texture features, and then changed the colour channel to calculate normalised and cumulative histograms in the wavelet domain on the MITOS-ATYPIA-14 dataset [111]. On the same dataset, Wan et al. applied a dual-tree complex wavelet transform (DT-CWT) to describe the images in the context of mitosis detection in breast cancer and the generalized Gaussian distribution and symmetric alpha-stable distribution parameters were used as features [108]. Tashk et al. combined features of LBP, morphometric, and statistical features extracted from mitotic candidates on the MITOS-12 dataset [112].

Recently, Mahmood et al. proposed a new methodology to imply post-processing techniques using feature extraction of HOG, LBP, statistical, and colour features to refine the detected mitosis cell as accepted or rejected through a threshold value based on the extracted features on the MITOS-ATYPIA datasets [59]. Bardou et al. developed two approaches where the first approach was an extracting local descriptors of dense scale invariant feature transform (DSIFT) features and speeded-up robust features (SURF) to be encoded by two coding models (bag of words and locality constrained linear coding) on the BreakHis dataset [26].

3.2.3. Feature Selection (Dimension Reduction)

Feature selection is the selection of a subset of the relevant features used in the model construction [113]. In machine learning, what we want is to avoid feature redundancy and the 'curse of dimensionality' problem. The 'curse of dimensionality' suggests that the training data have an exceptionally low density and lead to inability to promise an accurate estimation result, which defeats the purpose of training for a high accuracy classification model. This phenomenon will eventually impact the generalization performance in a negative way, for example, unstable estimation, overfitting issues, and local convergence; the large estimation error can easily compromise the prediction advantage provided by their greater representation power [114].

Therefore, this process is crucial because popular classification methods such as artificial neural network (ANN) and support vector machine (SVM), which are highly efficient for classification problems, tend to be sensitive to the dimensionality of data [115]. Additionally, data that consist of complex features reflect on quality-related issues such as the presence of noise, outliers, missing or duplicate data, and data that are biased or unrepresentative [116]. This process to reduce dimensionality can significantly eliminate irrelevant features, while the reduction in noise in machine learning contexts can produce a more robust learning model due to the association of fewer features [116].

One of the most common traditional approaches for this process is constructing new dimensions by mapping the original feature space into a new feature space with reduced dimensions. Common techniques used in this process are principal component analysis (PCA) and using a Pearson correlation matrix to construct a hierarchical representation of the data [115]. Other techniques include linear discriminant analysis (LDA), independent component analysis, and manifold learning. These techniques take the auto-covariance to solve the problem by transforming the high dimensional correlated feature set to a reduced feature set with lower dimensions.

However, these traditional approaches focus on choosing the most relevant features but disregard the fundamental interdependent structure of the features [117]. Recently, popular approaches have used heuristic search methods to select essential features from

original feature space by applying methods like genetic algorithm, simulated annealing, boosting, and particle swarm optimization. In [117] and [118], the authors proposed to apply a particle swarm optimizer (PSO) as the feature selection method in reducing the high dimensionality. In [119], the authors applied a genetic algorithm (GA) to select the best features and perfect parameter values of the machine learning classifiers. These recent proposed methods focus on the disregarded fundamental interdependent structure of the features from traditional features selection method. Tambasco Bruno et al. reduced their feature space by using an analysis of variance (ANOVA) [107].

3.2.4. Classification

The final set of features will then be fed as input to a classifier to estimate the breast cancer classes. The following are a few commonly applied classification methods.

- Nearest Neighbour: A non-parametric approach which falls under supervised learning widely used for both pattern recognition and classification applications [120]. The algorithm predicts each new point being input to the closest distance frame arrival point in the data; the calculation for distance varies but Euclidean distance is a common approach [121]. Let p and q be two datapoints of n -dimensions, then distance between x and y can be expressed by Euclidean distance shown in Equation (5).

$$D(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (5)$$

Then, the algorithm compares the distance between points and classifies it into different categories [121]. Kumar et al. applied a k -nearest neighbour classifier to classify cancer and non-cancerous biopsy breast images, and have suggested that this classifier performs the best among their studies [83]. Murtaza et al. experimented with six different machine learning classifiers and showed that the KNN algorithm performs the best [20].

- Support Vector Machine (SVM): Vapnik et al. proposed this method which works by mapping input information (feature vectors) to a higher dimensional space to obtain a hyperplane that can separate the labels/classes [122]. An optimal hyperplane can be obtained by maximizing the distances between support vectors (the data points closest to the boundary of the class) of two classes [123–125]. Recently, several research studies on breast cancer using histopathology images were performed by applying SVM classifiers [15,30,31,54]. Korkmaz and Poyraz proposed a classification framework focusing on minimum redundancy, maximum relevance feature selection, and least square SVM (LSSVM); their results claimed to be 100% accurate with only four false negatives for benign tumours in a three-class problem; however, no further evaluation was performed [126]. Chan and Tuszynski applied SVM classifier on their fractal features; their results achieved 97.9% F-score for magnification level $40\times$ on the BreakHis dataset [23]. Bardou et al. have also applied an approach of SVM to classify the images using handcrafted features [26].
- Artificial Neural Network (ANN): ANN is inspired by human perception that can models complex nonlinear functions. The basic architecture of ANN starts by receiving input data x_i , calculating each of the pieces of input information by multiplying to its corresponding weight w_{ij} , and obtaining a weighted output $f(x_j)$, with the support of a defined activation function until reaching the output layer. Figure 11 below demonstrates the basic structure of a single neuron in a feed-forward ANN [127]. Kassani et al. applied a multi-layer perceptron classifier on four different benchmark datasets and achieved the highest accuracy of 98.13% [19].

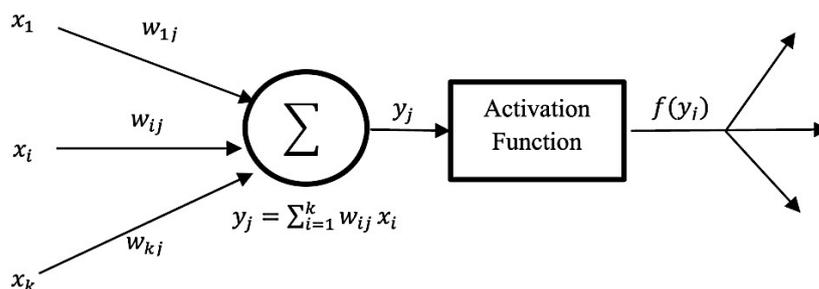


Figure 11. Basic structure of feed-forward ANN [127].

- Decision Tree: A decision tree algorithm is a supervised learning method for classification derived from the concept of ‘divide and conquer’ methodology. A complete decision tree is built based on feature space and labels; every new prediction will traverse from the root to the leaf node to produce an output. Asri et al. applied classification by using the C4.5 algorithm, an approach with a total of 11 features, and obtained 91.13% accuracy [128]. The extreme gradient boosting (XGBoost) is a new tree-based algorithm that has been increasing in popularity for data classification recently, and has proved to be a highly effective method for data classification [129]. Vo et al. have also applied gradient boosted trees as their breast cancer detection classifier [18].
- Bayesian Network: Bayesian network (BN) calculates probabilistic statistics to form a representation of relationships among a set of features space using an acrylic graph as shown in Figure 12, along with the value of conditional probabilities for each feature [130]. This type of classifier is commonly used for calculating probability estimations rather than predictions [116].

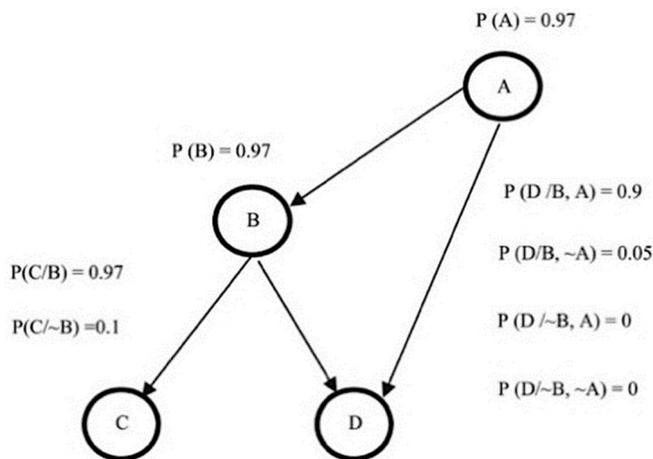


Figure 12. Bayesian network (BN) graph representation [130].

- Ensemble Classifier: This approach simply combines a few classifier methods instead of using a single classifier to produce a more accurate result. Commonly used methods to build an ensemble classifier are bagging, boosting, and random subspace method [131]. T.K. Ho proposed a random subspace classifier, in which a random feature subset is picked up from the original dataset for training each classifier; a voting scheme is then applied to produce a unique output from the from all the outputs in the combined classifiers [132]. Alkassar et al. applied an ensemble classifier that chooses the maximum score of prediction that includes a combination of decision tree, linear and quadratic discriminant, logistic regression, naive Bayes, SVM, and KNN [22].

3.3. Deep Learning CAD Methods

Following the recent advancements of deep learning (DL) that have shown a broad potential with state-of-the-art performance, many researchers have been approaching the process of feature extraction and selection using this automated technique. This improved approach combines learning and decision making by applying unsupervised learning upon different deep neural network architecture designs. It combines learning the features in histopathology images and classifying the images in one high complex architecture model. This process is often referred to as a black box and it can be complex to understand how deep learning works, i.e., how did the model come to this decision and what was involved in the learning process.

The deep learning approach is based on convolutional neural networks (CNN) to enable a deeper level of exploration and broaden the capability of a model to perform classification on breast cancer histology images. They are able to build a complex level of non-linear mapping of input and output by utilising cascaded convolutional layers. They are considered as a unique type of neural network where instead of having weights for each input, the weights are shared and are convolved across the input as a moving window [133]. They are computational models that are composed of multiple processing layers to retrieve features from raw data with multilevel representations and hierarchical abstraction [8]. A typical CNN consists of convolutional layer, activation function, pooling layer, and output layer. An example of a standard CNN model architecture with two feature stages is shown below in Figure 13 [134].

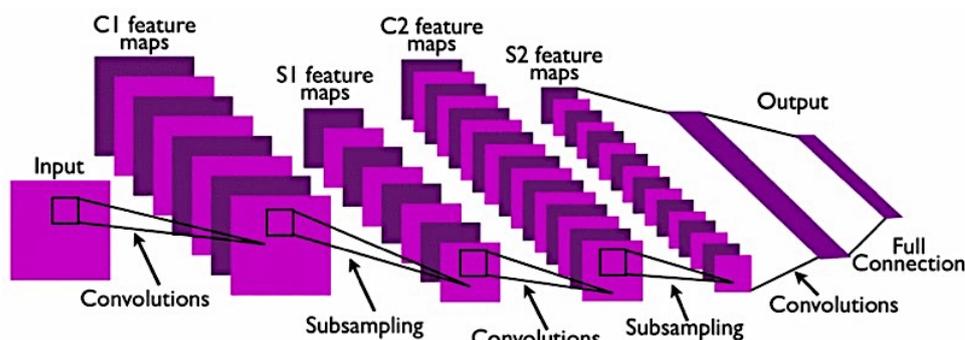


Figure 13. Basic structure of a CNN architecture [134].

To simplify, convolution is a signal processing operation which easily computes as a discrete spatial processing operation [121]. Recently, there have been several popular deep-learning-based models that improved the CNN model, such as AlexNet [135], VGGNet [136], GoogLeNet [137], Inception [138], DenseNet [139], Xception [140], and ResNet [141]. There are two ways to implement the method: (1) training from scratch and (2) transfer learning.

- **Training from scratch:** This method requires a large amount of input on histopathology images of breast cancer to train the CNN model. It requires more effort and skills to achieve a reliable performance CNN model when it comes to selecting hyperparameters such as learning rate, number of layers, convolutional filters and more, which can be a challenging task. This implementation also requires a high GPU processing power to perform training as CNN training can be time consuming because of the complex architecture [142].
- **Transfer learning:** Most publicly available datasets for breast histology images are considered as small datasets for training a deep learning model, which can be highly prone to overfitting due to the inferior performance of generalizability. The transfer learning method provides a solution to this by performing transfer knowledge tasks on the model based on a source domain that provides a large amount of sample data to the target domain. Pre-trained models can sufficiently prepare the small-scale histology dataset in a deep learning model. It can be used to: (1) perform as a

baseline model, which uses the architecture of the pre-trained network and builds the model from scratch by random initialization of weights [143]; (2) perform as a feature extractor, which extracts key features and the outputs which go into the convolutional base are fed directly to the classifier without modifying any weights or convolutional parameters [143]; and (3) perform fine tuning where weights will be passed into the designed network from the pre-trained network by fine tuning the layer or performing partial training of the network [143]. Figure 14 illustrates the transfer learning approach.

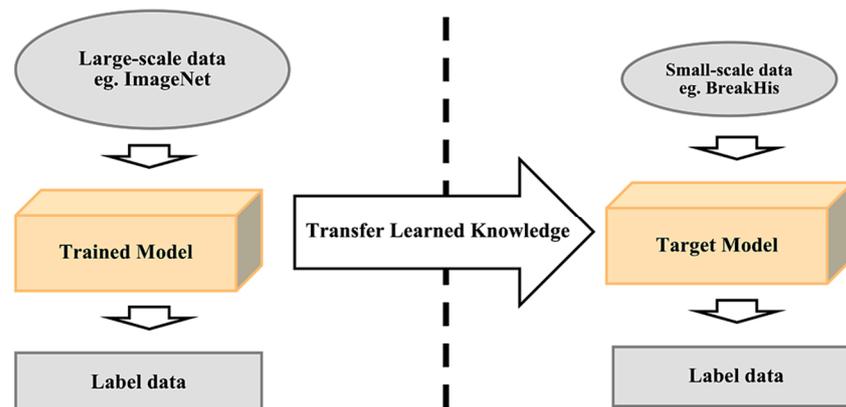


Figure 14. Transfer learning approach.

Bayramoglu et al. proposed two different CNN architectures: single-task CNN is used to predict malignancy and multi-task CNN is used to predict both malignancy and image magnification level simultaneously [17]. Gandomkar et al. proposed a two-step classification in which they first used a deep residual network (ResNet) with 152 layers trained for classifying patches from the images as benign or malignant for each magnification factor [27]. Then, they used the same pre-trained model to further classify the breast cancer sub-classes. Han et al. proposed a method class structure-based deep convolutional neural network (CSDCNN) based on GoogLeNet for eight-class classification of breast histopathological slides and have shown that their accuracy was higher for fine-tuning in comparison with training from scratch [25]. Spanhol et al. also adopted AlexNet and achieved a better result than a machine learning model trained with hand-crafted textural descriptors [24]. Alom et al. proposed a binary and multi classification for breast cancer methods using the Inception Recurrent Residual Convolutional Neural Network (IRRCNN) model and achieved 99.05% (for binary) and 98.59% (for multi) classification [11].

Toğaçar et al. proposed a novel method called BreastNet using CNN model architecture that adopted a multi-layer perception (MLP) as classifier [21]. Mahmood et al. performed a score-level fusion of Resnet-50 and Densenet-201 for classification [59]. Bardou et al. experimented with a second approach to apply a CNN model, and their results showed that deep learning approaches outperformed handcrafted features [26]. Sudharshan et al. have also shown their record of achieving the highest patient recognition rate (Prr) using a multiple-instance learning-based convolutional neural network (CNN) [30]. The research by Rakhlin et al. applied ResNet-50, InceptionV3, and VGG-16 models for feature extraction and a gradient boosting tree as classifier in their proposed methodology [53]. Shallu and Mehra applied transfer learning and demonstrated that pre-trained CNNs are good substitutes for the CNNs trained from scratch for the diagnosis of breast cancer using histopathology [60]. This is because training a CNN from scratch might take a lot more time, complexity, and effort to fine tune the model, especially if it has limited numbers of samples to train, whereas a pre-trained model does not suffer from this limitation. With this, a pre-trained CNN on the ImageNet [135] database provides a larger sample to feed into a CNN model to extract features more accurately and efficiently on histopathological images [144].

Cai et al. adopted modified faster-RCNN (regional convolutional neural network) for detecting mitosis cells using the Resnet-101 network pre-trained on ImageNet database to extract features for classification [145]. Mahmood et al. has also adopted the region-based CNN technique named Faster R-CNN [146] to perform detection of mitotic cells in breast cancer histology images [59]. The general architecture of Faster R-CNN consists of an extracted feature map from input image, followed by the generation of region proposal network (RPN) [147] and a classification network of deep CNN that detects the final mitotic cells as output. Vo et al. proposed a model called Inception-ResNet-v2 that combines CNNs of Inception and ResNet to train and extract visual features from multi-scale images to achieve both global and local features from breast tumours and feed them into a gradient boosting classifier [18]. George et al. proposed an approach for breast cancer diagnosis, which extracts features from nuclei based on a pre-trained set of CNN, namely, AlexNet, ResNet-18, and ResNet-50, on random patches obtained from histology images and finally classifies them with a SVM classifier [29]. Another study by Spanhol et al. proposed a method that combines a modified AlexNet and DeCAF [148] (or deep) features extraction that is based on reusing a previously trained CNN only as feature vectors, which is then used as input for a classifier trained only for the new classification task [149]. A method named Biopsy Microscopic Image Cancer Network (BMIC_Net) by Murtaza et al. has applied pre-trained AlexNet as feature extractor [20].

Budak et al. proposed a novel method that uses a fully convolutional network (FCN) transform from AlexNet as an encoder for high-level feature extraction; the output of the FCN will then be transformed to a one-dimensional sequence for classification using Bi-LSTM [28]. A recent model named Long Short-Term Memory (LSTM) [150] based on a recurrent neural network (RNN) was introduced and has increased in popularity due to its powerful ability. The authors adopted this architecture that combines a bidirectional RNN (Bi-RNN) that handles two sources of information and LSTM for classifying breast cancer [28]. Alkassar et al. used the Xception and DeseNet to perform extraction on shallow and deep features from breast histology images [22]. Araujo et al. combined a CNN model to extract features and a SVM classifier to perform breast cancer classification [31]. One of the most promising developed deep learning models was the lymph node assistant (LYNA) algorithm based on Inception-v3 by the researchers of Naval Medical Center San Diego and Google AI [55]. They adopted the Inception-v3 network because this model has been shown to achieve greater than 78.1% accuracy on Stanford's ImageNet dataset. Their results have successfully achieved a receiver operating characteristic area under curve (AUC) of 99% and a tumour-level sensitivity of 91% at 1 false positive per patient [55].

3.4. Performance Evaluation

To demonstrate the effectiveness of a CAD system for breast cancer diagnosis, it is important that we can evaluate our approaches to understand the performance of the system quantitatively as well as inspecting the underlying problems to be improved. In medical imaging diagnostic tests, sensitivity and specificity parameters are widely used to evaluate performance [40]. Other commonly used metrics for diagnosis evaluation are F1-measure [19,23,59,84,148,151], precision [19,59,60,148], accuracy [11,18–22,24–29,31,53,60,83,128,130,131] and receiver operating characteristics (ROC). The parameters are mostly calculated depending on the true positive (TP), true negative (TN), false positive (FP), and false negative (FN), which are the numbers of pixels corresponding to the parameters. To explain these variables, the true positive (TP) implies the number of patients who are predicted to be suffering cancer and are suffering. True negative (TN) implies the number of patients predicted to be not suffering cancer and in fact they are not suffering. False positive (FP) implies the number of patients who are predicted as cancer patients but in fact they are not suffering from cancer. False negative (FN) is the number of patients predicted as not cancer patients but in fact, they are suffering from cancer [40]. The following shows each parameter used for evaluating the performance of a classification model and its calculation formula [106].

- *Sensitivity* represents the percentage of positive numbers of samples classified correctly. The formula to calculate this is shown in Equation (6).

$$\text{Sensitivity (\%)} = \frac{TP}{TP + FN} \times 100 \quad (6)$$

- *Specificity* represents the percentage of negative numbers of samples classified correctly. The formula to calculate this is shown in Equation (7).

$$\text{Specificity (\%)} = \frac{TN}{TN + FP} \times 100 \quad (7)$$

- *Accuracy* represents the percentage of correct classification rate. The formula to calculate this is shown in Equation (8).

$$\text{Accuracy(\%)} = \frac{TP + TN}{\text{Number of Samples}} \times 100 \quad (8)$$

- *Precision* also known as PPV (Positive Predictive Value) represents the statistical variability measurement (total number of positive results). The formula to calculate this is shown in Equation (9).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

- *Recall* represents the proportion of negative numbers of samples classified correctly. The formula to calculate this is shown in Equation (10).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

- *F1-measure* represents the weighted mean of precision and recall. The formula to calculate this is shown in Equation (11).

$$F1 - \text{measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (11)$$

Besides the evaluation metrics, another useful technique in visualizing the performance of a classifier, specifically multiclassification, is by using receiver operating characteristics (ROC). This type of visual is represented in a two-dimensional graph to generate information on the trade-off of the true positive rate (sensitivity) and the false positive rate (1-specificity) within different thresholds. When evaluating a multi-class problem, each class generates different ROC graphs for comparison. To understand the ROC curve, we examine the area under the curve (AUC) to determine the capability of the features extracted for training a classifier. The larger area AUC indicates reliable performance of the model.

Recognition rate is also calculated to represent the multi-class performance on machine learning algorithms by measuring the patient-wise diagnosis [24]. The parameter is calculated as Equations (12) and (13).

$$\text{Patientscore}(P_s) = \frac{\text{Correctly classified cancer images of the patient}}{\text{Total number of cancer images of the patient}} \quad (12)$$

$$\text{Patient recognition rate}(Prr) = \frac{\sum P_s}{\text{Total number of patients}} \quad (13)$$

Table 6 demonstrates the comparison of different methods, datasets used, and evaluation results by different researchers for breast cancer diagnosis systems.

Table 6. Comparison of different methods for breast cancer diagnosis system.

Approach on CAD Method	Ref	Year	Dataset	Classification Type	Methods	Results
Conventional	[83]	2015	2828 histology images	Binary	KNN	Accuracy: 92.2% Specificity: 94.02% Sensitivity: 82% F1-measure: 75.94%
	[126]	2015	Firat University Medicine Faculty Pathology Laboratory	Multi-class	SVM (Least Square Support Vector Machine)	Accuracy: 100%; four FN for benign tumours in a three-class problem
	[15]	2016	BreaKHis	Multi-class	SVM, Random Forest, QDA (Quadratic Discriminant Analysis), Nearest Neighbour	Accuracy: 80% to 85%
	[23]	2016	BreaKHis	Binary and Multi-class	SVM	Highest F1-score: 97.9%
	[128]	2016	Wisconsin Breast Cancer dataset	Binary	Decision tree: C4.5 algorithm	Accuracy: 91.13%
	[128]	2016	Wisconsin Breast Cancer dataset	Binary	SVM	Accuracy: 97.13%
	[30]	2019	BreaKHis	Binary	SVM	Pr: 92.1%
	[22]	2021	BreaKHis	Binary and Multi-class	Ensemble Classifier	Highest accuracy: 99%

Table 6. Cont.

Approach on CAD Method	Ref	Year	Dataset	Classification Type	Methods	Results
Deep Learning	[26]	2018	BreakHis	Binary and Multi-class	SVM and CNN	Accuracy: 96.15–98.33% (binary); 83.31–88.23% (multi-class)
	[17]	2016	BreakHis	Binary and Multi-class	Single-task CNN (malignancy); Multi-task CNN (magnification level)	Prr: 83.13%. Prr: 80.10%
	[24]	2016	BreakHis	Multi-class	AlexNet CNN	Prr: 90%
	[30]	2019	BreakHis	Multi-class	MIL-CNN (Multiple Instance Learning-CNN)	Prr: 92.1%
	[53]	2018	BACH	Binary and Multi-class	ResNet-50, InceptionV3, VGG-16 and Gradient boosted trees	Accuracy: 87.2% (for binary) and 93.8% (for multi) AUC: 97.3% Sensitivity: 96.5 Specificity: 88.0%
	[60]	2018	BreakHis	Binary	VGG16, VGG19, ResNet5 and Logistic regression	Accuracy: 92.60% AUC: 95.65% Precision: 95.95%
	[149]	2017	BreakHis	Multi-class	Modified AlexNet and DeCAF (Deep Convolutional Activation Feature)	Accuracy: 81.5–86.3% F1-score: 86.7%-90.3%
	[31]	2017	Bioimaging Challenge 2015	Binary and Multi-class	CNN and SVM	Accuracy: 83.3%(binary); 77.8% (multi-class) Sensitivity: 95.6%.
	[25]	2017	BreakHis	Multi-class	Custom CSDCNN (Class Structure-based Deep Convolutional Neural Network) based on GoogLeNet	Accuracy: 93.2%
	[27]	2018	BreakHis	Binary and Multi-class	ResNet CNN	Accuracy: 98.77% (Binary) Prr: 96.25% (Multi class)
	[145]	2019	MITOS-ATYPIA-14, TUPAC-16	Binary	Modified faster-RCNN	Precision: 76% Recall: 72% F1 score: 73.6%
	[18]	2019	BreakHisBioimaging Challenge 2015	Multi-class	Inception and ResNet CNN (IRRCNN) and Gradient boosting trees	Accuracy: 99.5% (binary); 96.4% (multi-class)

Table 6. Cont.

Approach on CAD Method	Ref	Year	Dataset	Classification Type	Methods	Results
Deep Learning	[29]	2019	BreakHis	Multi-class	Pre-trained CNN (AlexNet, ResNet-18 and ResNet-50) and SVM	Accuracy: 96.88% Sensitivity: 97.30% Specificity: 95.97% AUC: 0.9942
	[54]	2019	Bioimaging Challenge 2015	Multi-class	Pre-trained ResNet50 with SVM classifier	Accuracy: 95% Recall: 89%
	[28]	2019	BreakHis	Binary	FCN (Fully Convolutional Network) based on AlexNet and Bi-LSTM (Bidirectional Long Short-Term Memory)	Accuracy: 91.90% Sensitivity: 96.8% Specificity: 91%
	[11]	2019	BreakHis and Bioimaging Challenge 2015	Binary and Multi-class	Inception Recurrent Residual CNN (IRRCNN)	Accuracy: 99.05% (for binary) and 98.59% (for multi)
	[55]	2019	Camelyon16	Binary	LYNA algorithm based on Inception-v3	AUC: 99% Sensitivity: 91%
	[19]	2020	BACH, BreakHis, PatchCamelyon, and Bioimaging 2015	Binary	Pre-trained VGG19, MobileNet, and DenseNet with MLP (Multi-Layer Perceptron)	Accuracy: 92.71% Precision: 95.74% Recall: 89.80% F-score: 92.43%
	[21]	2020	BreakHis	Multi-class	CNN features with MLP (Multi-Layer Perceptron)	Accuracy: 98.80%
	[59]	2020	MITOS-12MITOS-ATYPIA-14	Multi-class	Faster-RCNN and a score-level fusion of Resnet-50 and Densenet-201 CNNs	Precision: 87.6% Recall: 84.1% F1-measure: 85.8%
[20]	2020	BreakHis	Multi-class	BMIC_Net: Pre-trained AlexNet and KNN	Accuracy: 95.48%	
[22]	2021	BreakHis	Binary and Multi-class	Xception and DenseNet CNNs	Accuracy: 99% (binary); 92% (multi-class)	

4. Discussion and Conclusions

In this review paper, various techniques and approaches applied in every process of a CAD system have been discussed. In this section, we will discuss the importance of each process in developing a CAD system, the impact of the CAD system, and applicability of the system in the real world. To build a reliable CAD system, histopathology images that are being fed into a machine learning model perform better when they are refined and normalised. However, this process should not demolish any key features and biological tissues contained in the histopathology images. Diverse quality and intensities in images will directly affect how a machine learning model learns. Therefore, the consistency and accuracy of a machine learning CAD system depends highly on image qualities.

When it comes to segmentation, one of the critical conditions is to consider the problem of overlapping and obstructed boundaries in a histopathology image. A good segmentation technique will resolve this issue without demolishing any geometrical features. Based on our review, we can conclude that machine learning approaches perform better in tackling these issues in segmentation, for example using clustering, energy optimization, and feature-based techniques. As discussed earlier, the most common descriptive features extracted from histopathology images are morphological features, textural features, and graph topological features. Then, feature selection is applied to reduce the dimensionality of the set of features extracted to only select and focus on features that produce high impacts on the classification task. However, another approach is to use CNN topology to automatically learn the features from a histopathology image which is considered as a deep learning approach for feature extraction and selection. This approach for feature learning has been observed to perform better than manual descriptive features.

Based on this review study, the classifiers with the highest accuracy are SVM, ANN (CNN), and ensemble learning with ranges from 97.13–99% accuracy. Of course, the previous steps and design methodology of each of these approaches vary, but these classifiers are the ones that significantly contributed to the final performance. SVM models are highly efficient for non-linear decision boundaries with various kernels to select from. They are also high robustness techniques to tackle overfitting issues and dealing with high dimensional space. CNN derived from the ANN category are autonomous solutions that learn and gather information and knowledge from the images to make a decision. They are a powerful tool in analysing and processing data from grid-like topology [8] which includes images at pixel level. Ensemble learning is a robust approach as it uses multiple classifiers instead of one. It solves the issues of bias and invariance in classification task, which provides a more reliable output considering multiple factors.

The achievement and contribution of a CAD system impacts pathologists in examining breast cancer in several ways. As the process to examine a histology image requires time and effort under the microscope, it can be challenging as this type of visual contains highly complex patterns for a human eye to examine. Each process of a CAD system not only assists visualisation for pathologists, but it also provides a verification on each decision made. The impacts are discussed as follows:

- Image enhancement: Original histology images may contain visuals like noise, colour variation, intensity variation, low pixelation or more because of the staining processing during image acquisition. It is challenging to focus on the target area; therefore, image processing plays a role in standardising and improving the quality of histology images.
- Detecting the cells or nuclei: Segmentation procedure assists in locating and identifying every cell in the image. This plays a role in obtaining the accurate region of interest to further measure the existence of cancer in the cell.
- Learning the features: This process of feature extraction provides the geometrical information of the detected cell which will be later considered as knowledge to determine the possibility of cancer. The CNN approach on this matter provides a robust solution with automated learning.
- Justification on diagnosis results: There always exists a situation where pathologists might examine an incorrect result due to several factors such as lack of experience,

heavy workload, human error, or miscalculation. Thus, a CAD system can provide a second perspective on or verification of the diagnosis results by pathologists under the microscope assessment.

- Fast diagnosis results: As discussed in this paper, one of the benefits of a CAD system is to help breast cancer patients in early diagnosis to treat it before it progresses to more advanced stages. Pathologists often face challenges while diagnosing breast cancer because it requires an extensive amount time, effort, and process to perform microscope examination on histology images, therefore a CAD system can efficiently provide a faster solution.
- Improve productivity: The advancement of machine learning techniques produces higher productivity in a pathologist's microscope examination and possible reduction of the number of false negatives associated with morphologic detection of tumour cells for deep learning techniques [55].

Although CAD systems provide an optimistic aspect in assisting the medical image analysis and better performance, there exist several research gaps on using a CAD system in a real-world environment. Due to these limitations, it is challenging to adopt these tools in a real-world environment.

- Data limitations: Working with complex and large amounts of medical data can be challenging as they require high processing power and huge memory storage. Machine learning, especially deep learning, requires a large amount of data to train the model to produce a reliable and correct result. Some of the research papers acquired small datasets from private institutions, which are more likely to perform differently when being used in the real-world hospital environment. For publicly available datasets, most of them are considered as small datasets which are also most likely not applicable when it comes to performing in a real-world environment. Looking at the largest public dataset, for example the BreakHis dataset, it does not satisfy the condition of a dataset with enough patient samples. Therefore, existing CAD systems do not have sufficient knowledge learned that is ready to be applicable in the real-world environment.
- Bias and imbalance class: This problem among datasets can lead to undesired classification for the diagnosis result. When a CAD system is built upon a dataset with imbalanced classes, the results will be more likely to be biased and therefore produce wrong diagnosis. When a trained model is biased to a specific class due to the imbalanced dataset it destroys the reliability of a CAD system because it will increase the rate of wrong classification. There are solutions to deal with problems like these by applying oversampling, undersampling, and algorithm-level methods [152]. Therefore, there are insufficient investigations performed on solutions that show significant improvement for imbalanced data to be able to practically use it in hospitals.

This paper has presented the detailed process of designing a machine learning computer-aided diagnosis expert system for breast cancer on histopathology images using both conventional and deep learning approaches. Publicly available histopathology images have also captivated the interest of many developers and researchers in exploring the possibilities in the datasets. Machine-learning-based CAD systems have contributed a promising performance when compared to a diagnosis performed by a pathologist using a microscope. The advancement of deep learning has also remarkably outperformed the conventional approach on feature learning and capability of a CAD system. Analysing medical datasets of breast histopathological images is a challenging task due to differences and artifacts during image acquisition and because of the complex images. Therefore, techniques developed for analysing breast histology images require robustness to overcome all underlying variations. This review paper has explored the most recent developments in breast cancer diagnosis systems and provides a comparison overview of accuracy, benefits, disadvantages, and techniques employed by different researchers. A general review has been presented on techniques applied for classifying breast cancer, existing challenges, and the future direction of computer-aided diagnosis (CAD) systems for breast cancer.

5. Future Directions

Although many research studies presented in this paper have showed particularly reliable performance, there are still further possibilities to be explored in the future scope to further improve breast cancer diagnosis systems.

- Recently, the investigation and proposed CNN models have been increasing to provide an efficient solution to solve task-specific problems. In the future, there is always space for a new and more powerful CNN model that combines and utilises all the existing CNN's good characteristics to be discovered. For breast cancer classification, specifically a model that performs segmentation on cancer and non-cancerous regions.
- Most research studies focus on the indicators of accuracy and performance metrics while developing a diagnosis system. However, when it comes to the applicability in real-world hospital environments the performance is undefined. Problems like class imbalance and large-scale diagnosis systems require extensive investigation in unpredictable real-world environments to obtain reliable CAD systems. Therefore, further investigation needs to be performed and will require many years of clinical practice of a CAD system in the real-world environment to constantly adapt and improve to be able gain credibility for clinical adoption in the future.
- Currently, the development of pre-trained CNNs on histopathology breast cancer image datasets does not exist. Most of the current research studies apply feature extraction that uses pre-trained CNNs on the general ImageNet data. Therefore, future researchers can explore building a large-scale pre-trained CNN focusing on breast cancer histopathology images that is task specific to assist breast cancer diagnosis.
- In recent research studies, the authors in [26] have applied feature descriptors of scale invariant feature transform (DSIFT) features and speeded-up robust features (SURF). However, oriented fast and rotated brief (ORB) features have outperformed both SIFT and SURF [151]. In the future, further analysis can investigate the ORB features on a breast cancer classification task.
- It will be important to investigate a reliable-performance CAD system over a longer period with various settings to understand the strengths and weaknesses to ensure the confidence and reliability of the system to be integrated in practical healthcare in the future of medical diagnosis.
- Recently, new algorithms like eXtreme Gradient Boosting (XGBoost) [153] have shown increased popularity because of their reliable performance and can be experimented with and integrated in CAD systems.
- Developing a mobile-based compatible expert system for breast cancer diagnosis to provide further convenience for more users to access, especially those with limited access to computer-based systems.

Author Contributions: Conceptualization, X.Y.L.; writing—original draft preparation, X.Y.L.; writing—review and editing, X.Y.L., N.H. and J.C.; supervision, N.H. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in [15,31–33,35–39].

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

References

1. Ferlay, J.; Ervik, M.; Lam, F.; Colombet, M.; Mery, L.; Piñeros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Available online: <https://tinyurl.com/ugemlbs> (accessed on 16 March 2021).
2. Cancer Research UK Breast Cancer Statistics | Cancer Research UK. Available online: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero> (accessed on 16 March 2021).
3. Sizio, G.R.M.A.; Leite, C.R.M.; Guerreiro, A.M.G.; Neto, A.D.D. Fuzzy Method for Pre-Diagnosis of Breast Cancer from the Fine Needle Aspirate Analysis. *Biomed. Eng. Online* **2012**, *11*. [[CrossRef](#)]
4. Robertson, S.; Azizpour, H.; Smith, K.; Hartman, J. Digital Image Analysis in Breast Pathology—From Image Processing Techniques to Artificial Intelligence. *Transl. Res.* **2018**, *194*, 19–35. [[CrossRef](#)] [[PubMed](#)]
5. Krawczyk, B.; Schaefer, G.; Woźniak, M. A Hybrid Cost-Sensitive Ensemble for Imbalanced Breast Thermogram Classification. *Artif. Intell. Med.* **2015**, *65*. [[CrossRef](#)] [[PubMed](#)]
6. Bhardwaj, A.; Tiwari, A. Breast Cancer Diagnosis Using Genetically Optimized Neural Network Model. *Expert Syst. Appl.* **2015**, *42*. [[CrossRef](#)]
7. Chen, H.L.; Yang, B.; Liu, J.; Liu, D.Y. A Support Vector Machine Classifier with Rough Set-Based Feature Selection for Breast Cancer Diagnosis. *Expert Syst. Appl.* **2011**, *38*. [[CrossRef](#)]
8. Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
9. What Is Breast Cancer? | Cancer Research UK. Available online: <https://www.cancerresearchuk.org/about-cancer/breast-cancer/about> (accessed on 16 March 2021).
10. Breast Cancer Organization. What Is Breast Cancer? | Breastcancer.Org. 2016. pp. 1–19. Available online: https://www.breastcancer.org/symptoms/understand_bc/what_is_bc (accessed on 16 March 2021).
11. Alom, M.Z.; Yakopcic, C.; Nasrin, M.S.; Taha, T.M.; Asari, V.K. Breast Cancer Classification from Histopathological Images with Inception Recurrent Residual Convolutional Neural Network. *J. Digit. Imaging* **2019**, *32*. [[CrossRef](#)]
12. Akram, M.; Iqbal, M.; Daniyal, M.; Khan, A.U. Awareness and Current Knowledge of Breast Cancer. *Biol. Res.* **2017**, *50*, 33. [[CrossRef](#)]
13. Pantanowitz, L.; Evans, A.; Pfeifer, J.; Collins, L.; Valenstein, P.; Kaplan, K.; Wilbur, D.; Colgan, T. Review of the Current State of Whole Slide Imaging in Pathology. *J. Pathol. Inform.* **2011**, *2*. [[CrossRef](#)] [[PubMed](#)]
14. Khan, S.U.; Islam, N.; Jan, Z.; Ud Din, I.; Rodrigues, J.J.P.C. A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning. *Pattern Recognit. Lett.* **2019**, *125*. [[CrossRef](#)]
15. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Trans. Biomed. Eng.* **2016**, *63*. [[CrossRef](#)]
16. 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*. [[CrossRef](#)]
17. Bayramoglu, N.; Kannala, J.; Heikkila, J. Deep Learning for Magnification Independent Breast Cancer Histopathology Image Classification. In Proceedings of the International Conference on Pattern Recognition, Cancun, Mexico, 4–8 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 2440–2445. [[CrossRef](#)]
18. Vo, D.M.; Nguyen, N.Q.; Lee, S.W. Classification of Breast Cancer Histology Images Using Incremental Boosting Convolution Networks. *Inf. Sci.* **2019**, *482*. [[CrossRef](#)]
19. Kassani, S.H.; Kassani, P.H.; Wesolowski, M.J.; Schneider, K.A.; Deters, R. Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks. *arXiv* **2019**, arXiv:1909.11870.
20. Murtaza, G.; Shuib, L.; Mujtaba, G.; Raza, G. Breast Cancer Multi-Classification through Deep Neural Network and Hierarchical Classification Approach. *Multimed. Tools Appl.* **2020**, *79*. [[CrossRef](#)]
21. Toğaçar, M.; Özkurt, K.B.; Ergen, B.; Cömert, Z. BreastNet: A Novel Convolutional Neural Network Model through Histopathological Images for the Diagnosis of Breast Cancer. *Phys. A Stat. Mech. Its Appl.* **2020**, *545*. [[CrossRef](#)]
22. Alkassar, S.; Jebur, B.A.; Abdullah, M.A.M.; Al-Khalidy, J.H.; Chambers, J.A. Going Deeper: Magnification-Invariant Approach for Breast Cancer Classification Using Histopathological Images. *IET Comput. Vis.* **2021**, *15*, 151–164. [[CrossRef](#)]
23. Chan, A.; Tuszynski, J.A. Automatic Prediction of Tumour Malignancy in Breast Cancer with Fractal Dimension. *R. Soc. Open Sci.* **2016**, *3*. [[CrossRef](#)]
24. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. Breast Cancer Histopathological Image Classification Using Convolutional Neural Networks. In Proceedings of the International Joint Conference on Neural Networks, Vancouver, BC, Canada, 24–29 July 2016; IEEE: Piscataway, NJ, USA, 2016.
25. Han, Z.; Wei, B.; Zheng, Y.; Yin, Y.; Li, K.; Li, S. Breast Cancer Multi-Classification from Histopathological Images with Structured Deep Learning Model. *Sci. Rep.* **2017**, *7*. [[CrossRef](#)] [[PubMed](#)]
26. Bardou, D.; Zhang, K.; Ahmad, S.M. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. *IEEE Access* **2018**, *6*. [[CrossRef](#)]
27. Gandomkar, Z.; Brennan, P.C.; Mello-Thoms, C. MuDeRN: Multi-Category Classification of Breast Histopathological Image Using Deep Residual Networks. *Artif. Intell. Med.* **2018**, *88*. [[CrossRef](#)]
28. Budak, Ü.; Cömert, Z.; Rashid, Z.N.; Şengür, A.; Çibuk, M. Computer-Aided Diagnosis System Combining FCN and Bi-LSTM Model for Efficient Breast Cancer Detection from Histopathological Images. *Appl. Soft Comput. J.* **2019**, *85*. [[CrossRef](#)]

29. George, K.; Faziludeen, S.; Sankaran, P.; Paul, J.K. Deep Learned Nucleus Features for Breast Cancer Histopathological Image Analysis Based on Belief Theoretical Classifier Fusion. In Proceedings of the IEEE Region 10 Annual International Conference, Proceedings/TENCON, Kochi, India, 17–20 October 2019; IEEE: Piscataway, NJ, USA, 2019.
30. Sudharshan, P.J.; Petitjean, C.; Spanhol, F.; Oliveira, L.E.; Heutte, L.; Honeine, P. Multiple Instance Learning for Histopathological Breast Cancer Image Classification. *Expert Syst. Appl.* **2019**, *117*. [[CrossRef](#)]
31. Araujo, T.; Aresta, G.; Castro, E.; Rouco, J.; Aguiar, P.; Eloy, C.; Polonia, A.; Campilho, A. Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. *PLoS ONE* **2017**, *12*. [[CrossRef](#)]
32. Aresta, G.; Araújo, T.; Kwok, S.; Chennamsetty, S.S.; Safwan, M.; Alex, V.; Marami, B.; Prastawa, M.; Chan, M.; Donovan, M.; et al. BACH: Grand Challenge on Breast Cancer Histology Images. *Med. Image Anal.* **2019**, *56*. [[CrossRef](#)] [[PubMed](#)]
33. Litjens, G.; Bandi, P.; Bejnordi, B.E.; Geessink, O.; Balkenhol, M.; Bult, P.; Halilovic, A.; Hermsen, M.; van de Loo, R.; Vogels, R.; et al. 1399 H&E-Stained Sentinel Lymph Node Sections of Breast Cancer Patients: The CAMELYON Dataset. *Gigascience* **2018**, *7*, giy065.
34. Bándi, P.; Geessink, O.; Manson, Q.; Van Dijk, M.; Balkenhol, M.; Hermsen, M.; Ehteshami Bejnordi, B.; Lee, B.; Paeng, K.; Zhong, A.; et al. From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The CAMELYON17 Challenge. *IEEE Trans. Med. Imaging* **2019**, *38*. [[CrossRef](#)]
35. Veeling, B.S.; Linmans, J.; Winkens, J.; Cohen, T.; Welling, M. Rotation Equivariant CNNs for Digital Pathology. In Proceedings of the 21st International Conference, Granada, Spain, 16–20 September 2018; Springer: Cham, Switzerland, 2018. Lecture Notes in Computer Science. Volume 11071.
36. Roux, L.; Racoceanu, D.; Loménie, N.; Kulikova, M.; Irshad, H.; Klossa, J.; Capron, F.; Genestie, C.; Naour, G.; Gurcan, M. Mitosis Detection in Breast Cancer Histological Images an ICPR 2012 Contest. *J. Pathol. Inform.* **2013**, *4*. [[CrossRef](#)]
37. MITOS-ATYPIA-14 Grand Challenge. Available online: <https://mitos-atypia-14.grand-challenge.org/> (accessed on 17 March 2021).
38. Veta, M.; Heng, Y.J.; Stathonikos, N.; Bejnordi, B.E.; Beca, F.; Wollmann, T.; Rohr, K.; Shah, M.A.; Wang, D.; Rousson, M.; et al. Predicting Breast Tumor Proliferation from Whole-Slide Images: The TUPAC16 Challenge. *Med. Image Anal.* **2019**, *54*. [[CrossRef](#)]
39. Drelie Gelasca, E.; Obara, B.; Fedorov, D.; Kvilekval, K.; Manjunath, B.S. A Biosegmentation Benchmark for Evaluation of Bioimage Analysis Methods. *BMC Bioinform.* **2009**, *10*. [[CrossRef](#)]
40. Kaushal, C.; Bhat, S.; Koundal, D.; Singla, A. Recent Trends in Computer Assisted Diagnosis (CAD) System for Breast Cancer Diagnosis Using Histopathological Images. *IRBM* **2019**, *40*, 211–227. [[CrossRef](#)]
41. De Matos, J.; De Souza Britto, A., Jr.; Oliveira, L.E.S.; Koerich, A.L. Histopathologic Image Processing: A Review. *arXiv* **2019**, arXiv:1904.07900.
42. Ciompi, F.; Geessink, O.; Bejnordi, B.E.; De Souza, G.S.; Baidoshvili, A.; Litjens, G.; Van Ginneken, B.; Nagtegaal, I.; Van Der Laak, J. The Importance of Stain Normalization in Colorectal Tissue Classification with Convolutional Networks. In Proceedings of the International Symposium on Biomedical Imaging, Melbourne, Australia, 18–21 April 2017; IEEE: Piscataway, NJ, USA, 2017.
43. Reinhard, E.; Ashikhmin, M.; Gooch, B.; Shirley, P. Color Transfer between Images. *IEEE Comput. Graph. Appl.* **2001**, *21*. [[CrossRef](#)]
44. Helmenstine, A.M. *Beer's Law Definition and Equation*; ThoughtCo.: New York, NY, USA, 2019.
45. Ruifrok, A.C.; Johnston, D.A. Quantification of Histochemical Staining by Color Deconvolution. *Anal. Quant. Cytol. Histol.* **2001**, *23*, 291–299.
46. 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*. [[CrossRef](#)] [[PubMed](#)]
47. Roy, S.; Kumar Jain, A.; Lal, S.; Kini, J. A Study about Color Normalization Methods for Histopathology Images. *Micron* **2018**, *114*, 42–61. [[CrossRef](#)] [[PubMed](#)]
48. Macenko, M.; Niethammer, M.; Marron, J.S.; Borland, D.; Woosley, J.T.; Guan, X.; Schmitt, C.; Thomas, N.E. A Method for Normalizing Histology Slides for Quantitative Analysis. In Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009, Boston, MA, USA, 28 June–1 July 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 1107–1110.
49. Kothari, S.; Phan, J.H.; Moffitt, R.A.; Stokes, T.H.; Hassberger, S.E.; Chaudry, Q.; Young, A.N.; Wang, M.D. Automatic Batch-Invariant Color Segmentation of Histological Cancer Images. In Proceedings of the International Symposium on Biomedical Imaging, Chicago, IL, USA, 30 March–2 April 2011; IEEE: Piscataway, NJ, USA, 2011.
50. Bejnordi, B.E.; Litjens, G.; Timofeeva, N.; Otte-Höller, I.; Homeyer, A.; Karssemeijer, N.; Van Der Laak, J.A.W.M. Stain Specific Standardization of Whole-Slide Histopathological Images. *IEEE Trans. Med. Imaging* **2016**, *35*. [[CrossRef](#)]
51. Sandid, F.; Douik, A. Texture Descriptor Based on Local Combination Adaptive Ternary Pattern. *IET Image Process.* **2015**, *9*. [[CrossRef](#)]
52. Van Der Laak, J.A.W.M.; Pahlplatz, M.M.M.; Hanselaar, A.G.J.M.; De Wilde, P.C.M. Hue-Saturation-Density (HSD) Model for Stain Recognition in Digital Images from Transmitted Light Microscopy. *Cytometry* **2000**, *39*. [[CrossRef](#)]
53. Rakhlin, A.; Shvets, A.; Iglovikov, V.; Kalinin, A.A. Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. In Proceedings of the 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, 27–29 June 2018; IEEE: Piscataway, NJ, USA, 2018; Volume 10882 LNCS.
54. Li, Y.; Wu, J.; Wu, Q. Classification of Breast Cancer Histology Images Using Multi-Size and Discriminative Patches Based on Deep Learning. *IEEE Access* **2019**, *7*. [[CrossRef](#)]

55. Liu, Y.; Kohlberger, T.; Norouzi, M.; Dahl, G.E.; Smith, J.L.; Mohtashamian, A.; Olson, N.; Peng, L.H.; Hipp, J.D.; Stumpe, M.C. Artificial Intelligence–Based Breast Cancer Nodal Metastasis Detection Insights into the Black Box for Pathologists. *Arch. Pathol. Lab. Med.* **2019**, *143*. [[CrossRef](#)] [[PubMed](#)]
56. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*. [[CrossRef](#)]
57. Tellez, D.; Litjens, G.; Bándi, P.; Bulten, W.; Bokhorst, J.M.; Ciompi, F.; van der Laak, J. Quantifying the Effects of Data Augmentation and Stain Color Normalization in Convolutional Neural Networks for Computational Pathology. *Med. Image Anal.* **2019**, *58*. [[CrossRef](#)] [[PubMed](#)]
58. Saxena, S.; Gyanchandani, M. Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review. *J. Med. Imaging Radiat. Sci.* **2020**, *51*, 42–61. [[CrossRef](#)] [[PubMed](#)]
59. Mahmood, T.; Arsalan, M.; Owais, M.; Lee, M.B.; Park, K.R. Artificial Intelligence-Based Mitosis Detection in Breast Cancer Histopathology Images Using Faster R-CNN and Deep CNNs. *J. Clin. Med.* **2020**, *9*, 749. [[CrossRef](#)]
60. Mehra, R. Breast Cancer Histology Images Classification: Training from Scratch or Transfer Learning? *ICT Express* **2018**, *4*. [[CrossRef](#)]
61. Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang, J. Color Image Segmentation: Advances and Prospects. *Pattern Recognit.* **2001**, *34*. [[CrossRef](#)]
62. Rouhi, R.; Jafari, M.; Kasaei, S.; Keshavarzian, P. Benign and Malignant Breast Tumors Classification Based on Region Growing and CNN Segmentation. *Expert Syst. Appl.* **2015**, *42*. [[CrossRef](#)]
63. Rundo, L.; Militello, C.; Vitabile, S.; Casarino, C.; Russo, G.; Midiri, M.; Gilardi, M.C. Combining Split-and-Merge and Multi-Seed Region Growing Algorithms for Uterine Fibroid Segmentation in MRgFUS Treatments. *Med. Biol. Eng. Comput.* **2016**, *54*. [[CrossRef](#)] [[PubMed](#)]
64. Marr, D.; Hildreth, E. Theory of Edge Detection. *Proc. R. Soc. Lond. Biol. Sci.* **1980**, *207*. [[CrossRef](#)]
65. Lu, C.; Ji, M.; Ma, Z.; Mandal, M. Automated Image Analysis of Nuclear Atypia in High-Power Field Histopathological Image. *J. Microsc.* **2015**, *258*. [[CrossRef](#)]
66. Vincent, O.; Folorunso, O. A Descriptive Algorithm for Sobel Image Edge Detection. In Proceedings of the 2009 InSITE Conference, Macon, GA, USA, 12–15 June 2009; Informing Science Institute California: Santa Rosa, CA, USA, 2009.
67. Salman, N. Image Segmentation Based on Watershed and Edge Detection Techniques. *Int. Arab. J. Inf. Technol.* **2006**, *3*, 104–110. [[CrossRef](#)]
68. Prewitt, J. Object enhancement and extraction. In *Picture Processing and Psychopictorics*; Elsevier: Amsterdam, The Netherlands, 1970; Volume 10.
69. Stehfest, H. Algorithm 368: Numerical Inversion of Laplace Transforms [D5]. *Commun. ACM* **1970**, *13*. [[CrossRef](#)]
70. Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *8*. [[CrossRef](#)]
71. George, Y.M.; Zayed, H.H.; Roushdy, M.I.; Elbagoury, B.M. Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images. *IEEE Syst. J.* **2014**, *8*. [[CrossRef](#)]
72. Gonzalez, R.C.; Woods, R.E.; Masters, B.R. Digital Image Processing, Third Edition. *J. Biomed. Opt.* **2009**, *14*. [[CrossRef](#)]
73. Faridi, P.; Danyali, H.; Helfroush, M.S.; Jahromi, M.A. An Automatic System for Cell Nuclei Pleomorphism Segmentation in Histopathological Images of Breast Cancer. In Proceedings of the 2016 IEEE Signal Processing in Medicine and Biology Symposium, SPMB 2016, Philadelphia, PA, USA, 3 December 2016; IEEE: Piscataway, NJ, USA, 2017.
74. Otsu, N. Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst Man Cybern* **1979**, *9*. [[CrossRef](#)]
75. Zarella, M.D.; Garcia, F.U.; Breen, D.E. A Template Matching Model for Nuclear Segmentation in Digital Images of H&E Stained Slides. In Proceedings of the 9th International Conference on Bioinformatics and Biomedical Technology, Lisbon, Portugal, 14–16 May 2017.
76. Saha, M.; Agarwal, S.; Arun, I.; Ahmed, R.; Chatterjee, S.; Mitra, P.; Chakraborty, C. Histogram Based Thresholding for Automated Nucleus Segmentation Using Breast Imprint Cytology. In *Advancements of Medical Electronics*; Springer: New Delhi, India, 2015.
77. Moncayo, R.; Romo-Bucheli, D.; Romero, E. A Grading Strategy for Nuclear Pleomorphism in Histopathological Breast Cancer Images Using a Bag of Features (BOF). In Proceedings of the 20th Iberoamerican Congress, CIARP 2015, Montevideo, Uruguay, 9–12 November 2015; Volume 9423.
78. Khairuzzaman, A.K.M.; Chaudhury, S. Multilevel Thresholding Using Grey Wolf Optimizer for Image Segmentation. *Expert Syst. Appl.* **2017**, *86*. [[CrossRef](#)]
79. Sirinukunwattana, K.; Khan, A.M.; Rajpoot, N.M. Cell Words: Modelling the Visual Appearance of Cells in Histopathology Images. *Comput. Med. Imaging Graph.* **2015**, *42*. [[CrossRef](#)]
80. Rokach, L.; Maimon, O. Clustering Methods. In *Data Mining and Knowledge Discovery Handbook*; Maimon, O., Rokach, L., Eds.; Springer US: Boston, MA, USA, 2005; pp. 321–352. ISBN 978-0-387-25465-4.
81. De Carvalho, F.D.A.T.; Lechevallier, Y.; De Melo, F.M. Partitioning Hard Clustering Algorithms Based on Multiple Dissimilarity Matrices. *Pattern Recognit.* **2012**, *45*. [[CrossRef](#)]
82. Kowal, M.; Filipczuk, P.; Obuchowicz, A.; Korbicz, J.; Monczak, R. Computer-Aided Diagnosis of Breast Cancer Based on Fine Needle Biopsy Microscopic Images. *Comput. Biol. Med.* **2013**, *43*. [[CrossRef](#)]
83. Kumar, R.; Srivastava, R.; Srivastava, S. Detection and Classification of Cancer from Microscopic Biopsy Images Using Clinically Significant and Biologically Interpretable Features. *J. Med. Eng.* **2015**. [[CrossRef](#)] [[PubMed](#)]

84. Shi, P.; Zhong, J.; Huang, R.; Lin, J. Automated Quantitative Image Analysis of Hematoxylin-Eosin Staining Slides in Lymphoma Based on Hierarchical Kmeans Clustering. In Proceedings of the 2016 8th International Conference on Information Technology in Medicine and Education, ITME 2016, Fuzhou, China, 23–25 December 2016; IEEE: Piscataway, NJ, USA, 2017.
85. Maqlin, P.; Thamburaj, R.; Mammen, J.J.; Manipadam, M.T. Automated Nuclear Pleomorphism Scoring in Breast Cancer Histopathology Images Using Deep Neural Networks. In Proceedings of the Third International Conference, MIKE 2015, Hyderabad, India, 9–11 December 2015; Volume 9468.
86. Belsare, A.D.; Mushrif, M.M.; Pangarkar, M.A.; Meshram, N. Classification of Breast Cancer Histopathology Images Using Texture Feature Analysis. In *Proceedings of the IEEE Region 10 Annual International Conference, Proceedings/TENCON*; IEEE: Piscataway, NJ, USA, 2016.
87. Wan, T.; Cao, J.; Chen, J.; Qin, Z. Automated Grading of Breast Cancer Histopathology Using Cascaded Ensemble with Combination of Multi-Level Image Features. *Neurocomputing* **2017**, *229*. [[CrossRef](#)]
88. Zhang, W.; Li, H. Automated Segmentation of Overlapped Nuclei Using Concave Point Detection and Segment Grouping. *Pattern Recognit.* **2017**, *71*. [[CrossRef](#)]
89. Jia, D.; Zhang, C.; Wu, N.; Guo, Z.; Ge, H. Multi-Layer Segmentation Framework for Cell Nuclei Using Improved GVF Snake Model, Watershed, and Ellipse Fitting. *Biomed. Signal. Process. Control* **2021**, *67*, 102516. [[CrossRef](#)]
90. Janowczyk, A.; Madabhushi, A. Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases. *J. Pathol. Inform.* **2016**, *7*. [[CrossRef](#)] [[PubMed](#)]
91. Song, Y.; Zhang, L.; Chen, S.; Ni, D.; Lei, B.; Wang, T. Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning. *IEEE Trans. Biomed. Eng.* **2015**, *62*. [[CrossRef](#)] [[PubMed](#)]
92. Xu, Z.; Huang, J. Detecting 10,000 Cells in One Second. In Proceedings of the 19th International Conference, Athens, Greece, 17–21 October 2016; Volume 9901 LNCS.
93. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images. *Med. Image Anal.* **2019**, *58*. [[CrossRef](#)] [[PubMed](#)]
94. Zarella, M.D.; Breen, D.E.; Reza, M.A.; Milutinovic, A.; Garcia, F.U. Lymph Node Metastasis Status in Breast Carcinoma Can Be Predicted via Image Analysis of Tumor Histology. *Anal. Quant. Cytol. Histol.* **2015**, *37*, 273–285.
95. Belsare, A.D.; Mushrif, M.M.; Pangarkar, M.A.; Meshram, N. Breast Histopathology Image Segmentation Using Spatio-Colour-Texture Based Graph Partition Method. *J. Microsc.* **2016**, *262*. [[CrossRef](#)]
96. Krig, S. *Computer Vision Metrics: Survey, Taxonomy, and Analysis*; Springer: Berlin/Heidelberg, Germany, 2014.
97. Das, A.; Nair, M.S.; Peter, S.D. Computer-Aided Histopathological Image Analysis Techniques for Automated Nuclear Atypia Scoring of Breast Cancer: A Review. *J. Digit. Imaging* **2020**, *33*, 1091–1121. [[CrossRef](#)] [[PubMed](#)]
98. Zheng, Y.; Jiang, Z.; Xie, F.; Zhang, H.; Ma, Y.; Shi, H.; Zhao, Y. Feature Extraction from Histopathological Images Based on Nucleus-Guided Convolutional Neural Network for Breast Lesion Classification. *Pattern Recognition.* **2017**, *71*. [[CrossRef](#)]
99. Sharma, H.; Zerbe, N.; Lohmann, S.; Kayser, K.; Hellwich, O.; Hufnagel, P. A Review of Graph-Based Methods for Image Analysis in Digital Histopathology. *Diagn. Pathol.* **2015**, *1*. [[CrossRef](#)]
100. Shi, J.; Wu, J.; Li, Y.; Zhang, Q.; Ying, S. Histopathological Image Classification with Color Pattern Random Binary Hashing-Based PCANet and Matrix-Form Classifier. *IEEE J. Biomed. Health Inform.* **2017**, *21*. [[CrossRef](#)] [[PubMed](#)]
101. Ehteshami Bejnordi, B.; Lin, J.; Glass, B.; Mullooly, M.; Gierach, G.L.; Sherman, M.E.; Karssemeijer, N.; Van Der Laak, J.; Beck, A.H. Deep Learning-Based Assessment of Tumor-Associated Stroma for Diagnosing Breast Cancer in Histopathology Images. In Proceedings of the Proceedings International Symposium on Biomedical Imaging, Melbourne, Australia, 18–21 April 2017; IEEE: Piscataway, NJ, USA, 2017.
102. Zwanenburg, A.; Leger, S.; Vallières, M.; Löck, S. Image biomarker standardisation initiative-feature definitions. *arXiv* **2016**, arXiv:1612.07003.
103. Balazsi, M.; Blanco, P.; Zoroquiain, P.; Levine, M.D.; Burnier, M.N. Invasive Ductal Breast Carcinoma Detector That Is Robust to Image Magnification in Whole Digital Slides. *J. Med. Imaging* **2016**, *3*. [[CrossRef](#)]
104. Gupta, V.; Bhavsar, A. Breast Cancer Histopathological Image Classification: Is Magnification Important? In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017.
105. Rezaeilouyeh, H.; Mollahosseini, A.; Mahoor, M.H. Microscopic Medical Image Classification Framework via Deep Learning and Shearlet Transform. *J. Med. Imaging* **2016**, *3*. [[CrossRef](#)]
106. Shukla, K.K.; Tiwari, A.; Sharma, S. Classification of Histopathological Images of Breast Cancerous and Non Cancerous Cells Based on Morphological Features. *Biomed. Pharmacol. J.* **2017**, *10*. [[CrossRef](#)]
107. Tambasco Bruno, D.O.; Do Nascimento, M.Z.; Ramos, R.P.; Batista, V.R.; Neves, L.A.; Martins, A.S. LBP Operators on Curvlet Coefficients as an Algorithm to Describe Texture in Breast Cancer Tissues. *Expert Syst. Appl.* **2016**, *55*. [[CrossRef](#)]
108. Wan, T.; Zhang, W.; Zhu, M.; Chen, J.; Achim, A.; Qin, Z. Automated Mitosis Detection in Histopathology Based on Non-Gaussian Modeling of Complex Wavelet Coefficients. *Neurocomputing* **2017**, *237*. [[CrossRef](#)]
109. Gandomkar, Z.; Brennan, P.C.; Mello-Thoms, C. Computer-Assisted Nuclear Atypia Scoring of Breast Cancer: A Preliminary Study. *J. Digit. Imaging* **2019**, *32*. [[CrossRef](#)]
110. Khan, A.M.; Sirinukunwattana, K.; Rajpoot, N. A Global Covariance Descriptor for Nuclear Atypia Scoring in Breast Histopathology Images. *IEEE J. Biomed. Health Inform.* **2015**, *19*. [[CrossRef](#)] [[PubMed](#)]

111. Maroof, N.; Khan, A.; Qureshi, S.A.; Rehman, A.U.; Khalil, R.K.; Shim, S.O. Mitosis Detection in Breast Cancer Histopathology Images Using Hybrid Feature Space. *Photodiagn. Photodyn. Ther.* **2020**, *31*. [[CrossRef](#)]
112. Tashk, A.; Helfroush, M.S.; Danyali, H.; Akbarzadeh-jahromi, M. Automatic Detection of Breast Cancer Mitotic Cells Based on the Combination of Textural, Statistical and Innovative Mathematical Features. *Appl. Math. Model.* **2015**, *39*. [[CrossRef](#)]
113. Burges, C.J.C. Dimension Reduction: A Guided Tour. In *Foundation and Trends in Machine Learning*; now Publishers Inc.: Hanover, MA, USA, 2009; Volume 2. [[CrossRef](#)]
114. Clarke, R.; Ressom, H.W.; Wang, A.; Xuan, J.; Liu, M.C.; Gehan, E.A.; Wang, Y. The Properties of High-Dimensional Data Spaces: Implications for Exploring Gene and Protein Expression Data. *Nat. Rev. Cancer* **2008**, *8*, 37–49. [[CrossRef](#)] [[PubMed](#)]
115. Eisen, M.B.; Spellman, P.T.; Brown, P.O.; Botstein, D. Cluster Analysis and Display of Genome-Wide Expression Patterns. *Proc. Natl. Acad. Sci. USA* **1998**, *95*. [[CrossRef](#)]
116. Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine Learning Applications in Cancer Prognosis and Prediction. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 8–17. [[CrossRef](#)]
117. Sakri, S.B.; Abdul Rashid, N.B.; Muhammad Zain, Z. Particle Swarm Optimization Feature Selection for Breast Cancer Recurrence Prediction. *IEEE Access* **2018**, *6*. [[CrossRef](#)]
118. Cruz-Roa, A.; Gilmore, H.; Basavanhally, A.; Feldman, M.; Ganesan, S.; Shih, N.; Tomaszewski, J.; Madabhushi, A.; González, F. High-Throughput Adaptive Sampling for Whole-Slide Histopathology Image Analysis (HASHI) via Convolutional Neural Networks: Application to Invasive Breast Cancer Detection. *PLoS ONE* **2018**, *13*. [[CrossRef](#)]
119. Dhahri, H.; Al Maghayreh, E.; Mahmood, A.; Elkilani, W.; Faisal Nagi, M. Automated Breast Cancer Diagnosis Based on Machine Learning Algorithms. *J. Healthc. Eng.* **2019**. [[CrossRef](#)]
120. Rajaguru, H.; Sannasi Chakravarthy, S.R. Analysis of Decision Tree and K-Nearest Neighbor Algorithm in the Classification of Breast Cancer. *Asian Pac. J. Cancer Prev.* **2019**, *20*. [[CrossRef](#)]
121. Wadkar, K.; Pathak, P.; Wagh, N. Breast Cancer Detection Using ANN Network and Performance Analysis With SVM. *Int. J. Comput. Eng. Technol.* **2019**, *10*. [[CrossRef](#)]
122. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*. [[CrossRef](#)]
123. Introduction to Support Vector Machines—OpenCV 2.4.13.7 Documentation. Available online: https://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html (accessed on 17 March 2021).
124. Ayat, N.E.; Cheriet, M.; Suen, C.Y. Automatic Model Selection for the Optimization of SVM Kernels. *Pattern Recognit.* **2005**, *38*. [[CrossRef](#)]
125. Wang, H.; Zheng, B.; Yoon, S.W.; Ko, H.S. A Support Vector Machine-Based Ensemble Algorithm for Breast Cancer Diagnosis. *Eur. J. Oper. Res.* **2018**, *267*. [[CrossRef](#)]
126. Korkmaz, S.A.; Poyraz, M. Least Square Support Vector Machine and Minimum Redundancy Maximum Relevance for Diagnosis of Breast Cancer from Breast Microscopic Images. *Procedia Soc. Behav. Sci.* **2015**, *174*. [[CrossRef](#)]
127. Jha, G.K. *Artificial Neural Networks and Its Applications*; Indian Agricultural Statistics Research Institute (I.C.A.R.): New Delhi, India, 2007.
128. Asri, H.; Mousannif, H.; Al Moatassime, H.; Noel, T. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Comput. Sci.* **2016**, *83*, 1064–1069.
129. Parashar, J.; Rai, M. Breast Cancer Images Classification by Clustering of ROI and Mapping of Features by CNN with XGBOOST Learning. *Mater. Today Proc.* **2020**. [[CrossRef](#)]
130. Kotsiantis, S.B. Supervised Machine Learning: A Review of Classification Techniques. *Informatica* **2007**, *31*, 249–268.
131. Min, S.H. A Genetic Algorithm-Based Heterogeneous Random Subspace Ensemble Model for Bankruptcy Prediction. *Int. J. Appl. Eng. Res.* **2016**, *11*, 2927–2931.
132. Ho, T.K. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*. [[CrossRef](#)]
133. Wahab, N.; Khan, A.; Lee, Y.S. Two-Phase Deep Convolutional Neural Network for Reducing Class Skewness in Histopathological Images Based Breast Cancer Detection. *Comput. Biol. Med.* **2017**, *85*. [[CrossRef](#)]
134. LeCun, Y.; Kavukcuoglu, K.; Farabet, C. Convolutional Networks and Applications in Vision. In Proceedings of the ISCAS 2010—2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, Paris, France, 2 June–30 May 2010.
135. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [[CrossRef](#)]
136. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015.
137. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway, NJ, USA, 2015.
138. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 7–9 July 2015.

139. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017.
140. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017.
141. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016.
142. Bevilacqua, V.; Brunetti, A.; Guerriero, A.; Trotta, G.F.; Telegrafo, M.; Moschetta, M. A Performance Comparison between Shallow and Deeper Neural Networks Supervised Classification of Tomosynthesis Breast Lesions Images. *Cogn. Syst. Res.* **2019**, *53*, 3–19. [[CrossRef](#)]
143. Sharma, S.; Mehra, R. Conventional Machine Learning and Deep Learning Approach for Multi-Classification of Breast Cancer Histopathology Images—A Comparative Insight. *J. Digit. Imaging* **2020**, *33*, 632–654. [[CrossRef](#)]
144. Komura, D.; Ishikawa, S. Machine Learning Methods for Histopathological Image Analysis. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 34–42. [[CrossRef](#)] [[PubMed](#)]
145. Cai, D.; Sun, X.; Zhou, N.; Han, X.; Yao, J. Efficient Mitosis Detection in Breast Cancer Histology Images by RCNN. In Proceedings of the International Symposium on Biomedical Imaging, Venice, Italy, 8–11 April 2019.
146. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*. [[CrossRef](#)] [[PubMed](#)]
147. Zhong, Z.; Sun, L.; Huo, Q. An Anchor-Free Region Proposal Network for Faster R-CNN-Based Text Detection Approaches. *IJDAR* **2019**, *22*, 315–327. [[CrossRef](#)]
148. Donahue, J.; Jia, Y.; Vinyals, O.; Hoffman, J.; Zhang, N.; Tzeng, E.; Darrell, T. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In Proceedings of the 31st International Conference on Machine Learning, ICML 2014, Beijing, China, 21–26 June 2014.
149. Spanhol, F.A.; Cavalin, P.R.; Oliveira, L.S.; Petitjean, C.; Heutte, L. Deep Features for Breast Cancer Histopathological Image Classification. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017, Banff, AB, Canada, 5–8 October 2017.
150. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*. [[CrossRef](#)]
151. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An Efficient Alternative to SIFT or SURF. In Proceedings of the IEEE International Conference on Computer Vision 2011, Barcelona, Spain, 6–13 November 2011; IEEE: Piscataway, NJ, USA, 2011.
152. He, H.; Garcia, E.A. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*. [[CrossRef](#)]
153. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.