Chapter 2

Background and Related Works

2.1 Background

Recent advancements in AI technology have a large impact on the healthcare industry and have added a new wing to the innovation of healthcare services, which are now rapidly transitioning to smart healthcare. The enormous success of AI approaches in various fields inspired researchers and medical experts to design the CAD for analysis of critical diseases such as breast cancer. The inability to obtain annotated data and analysis reports on time is hampered by the absence of an expert pathologist. As a result, it is also a primary concern to reduce the burden on pathologists. In recent years, there has been a significant increase in the number of studies focusing on CAD of breast cancer. Our primary goal in this thesis is to provide effective and efficient AI solutions for HBC and CMT classification.

Therefore, this chapter provides a systematic review of studies published in the last ten years in the field of AI-based frameworks for breast cancer histopathology. A few early works that are over ten years old but have made significant contributions to the field are also included.

2.1.1 Breast cancer and CAD in cancer histopathology

Cancer continues to be the leading cause of death worldwide, accounting for approximately 10 million deaths in 2020 alone [15]. According to the WHO most recent cancer statistics, 19.29 million new cancer cases were recorded worldwide in 2020. Among these cancers, breast cancer is the most prevalent with 2.26 million cases, followed by lung cancer with 2.21 million cases. Breast cancer is the most common cancer in women, accounting for 685,000 deaths worldwide in 2020 [16]. Besides that, the WHO predicts that the number of new breast cancer patients will rise by 70% over the next two decades. The current late-stage survival rate for breast cancer is around 30%. Early and accurate diagnosis is critical in improving prognosis and increasing the survival rate of breast cancer patients from 30% to 50% [15]. In general, breast tumours are classified into two types: benign and malignant. Non-invasive tumours are benign, whereas invasive cancerous tumours are malignant. It is critical to first classify tumours as benign or malignant in order to select the best treatment plan. Mammography, ultrasound imaging or sonograms, magnetic resonance imaging, computed tomography, and histopathological image analysis are the techniques used to diagnose HBC [17, 18]. In histopathological image analysis, tumour tissues or biopsy samples from an abnormal breast region are collected and fixed on microscopic glass slides. The sections are then stained with H&E stain and examined under a microscope by a pathologist for detecting cancerous changes in tissues. Aside from breast cancer, histopathology imaging is the gold standard for many types of cancer, including liver, lung, and bladder cancer [19]. In cases where more than one pathologist is available, a final decision is made only after consensus between two pathologists; otherwise, findings are reported by one pathologist only.

Nonetheless, there are three major issues with manual histopathology image analysis [20]. Firstly, there is a scarcity of trained oncopathologists, and it is rare to find more than one expert pathologist in the same location. Secondly, the manual slide analysis procedure is time consuming, tedious, and laborious, requiring a skilled pathologist to examine slides under a microscope for hours. As a result, pathologists may become fatigued and lose concentration while analyzing images. Finally, a reliable breast cancer subtype identification is dependent on an expert pathologist's professional experience and domain knowledge. These issues may lead to a misdiagnosis, particularly in the early stages of breast cancer. However, computer-aided diagnosis (CAD) systems can be used as a second opinion to solve classification problems in breast cancer. A CAD system is an economical, readily available, quick, and dependable source of early diagnosis [21, 22]. This system assists radiologists and physicians in identifying abnormalities through the use of various imaging modalities, which has reduced mortality rates from 30% to 70% [23]. As a result, such a system reduces human dependency, increases diagnosis rates, and lowers overall treatment costs by reducing false positive and false negative predictions [24]. Furthermore, an increased false-negative rate may result in no treatment for a breast cancer carrier, and misdiagnoses are most common in the early stages of breast cancer. Sadaf et al. [22], also reported that using a CAD system for breast cancer classification improves sensitivity by 10%.

2.1.2 Convolutional neural network

Convolutional Neural Network (CNN) is one of the most widely used architectures for deep learning in computer vision. The overall CNN process is depicted in Figure 2.1. A standard CNN model consists of a series of layers, namely convolutional layers, pooling layers, and fully connected layers, which are described below.

1. Convolutional Layers: A convolutional layer is the core building block of CNNs and is responsible for performing most computationally intensive tasks. Its parameter consists of several learnable filters. These filters in the initial layer extract low-level features, while deeper layers extract high-level features. Further, these extracted features present a feature map used by the successive layer. Basically,

the strength of this layer lies in the fact that correlation among neighbouring pixels is learned by local connectives, and simultaneously the number of parameters are reduced by the weight sharing process in the same feature map. The convolution operation is depicted in Figure 2.2.

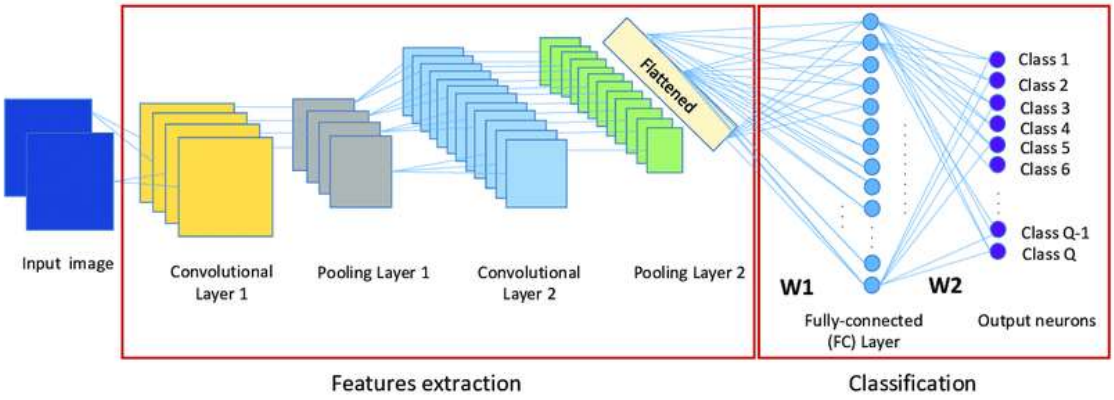

Figure 2.1: A general overview of the CNN architecture with two convolution layers and one fully-connected layers.

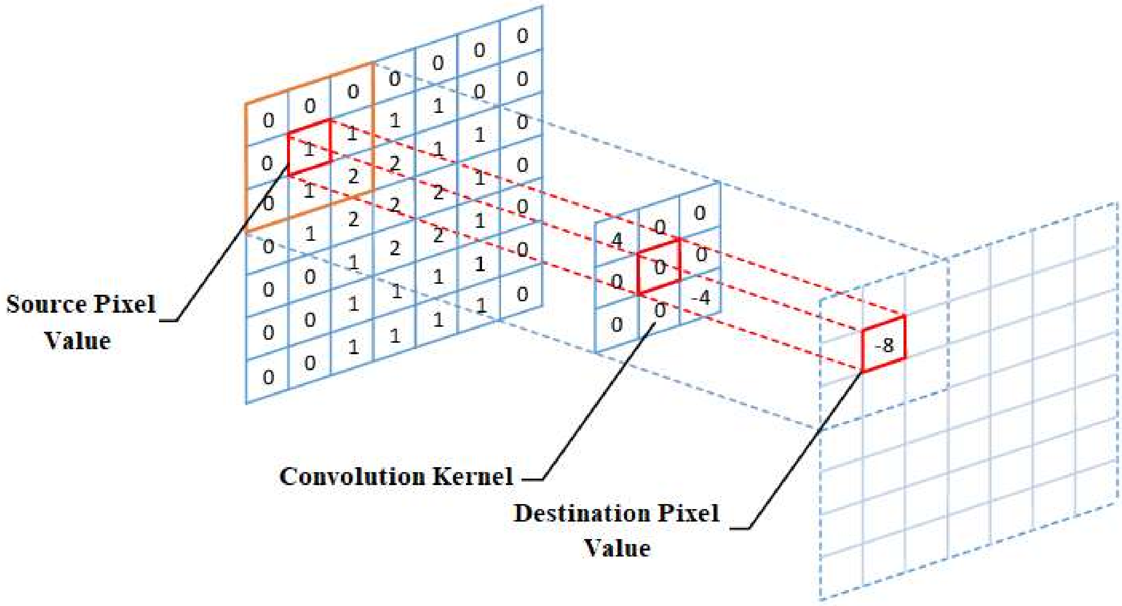

Figure 2.2: An example of a convolution operation on a 7x7 matrix with a 3x3 kernel

Mathematically convolutional layer is defined by following equation:

$$\psi^1 = g(W^1 * \psi^0 + b^1) \tag{2.1}$$

where $*$ is the convolution operator, $g(\cdot)$ represents activation function which may be ReLU (rectified linear unit), hyperbolic etc. The input image is denoted by ψ^0, b^1 is bias term, W^1 is filter and ψ^1 is output feature maps of convolution layer and input of pooling layer.

2. Pooling Layers: The pooling layer is periodically placed in between successive convolutional layers in a CNN framework. Its major objective is to minimize the computation and amount of network parameters by progressively reducing the dimension of feature maps. Average-pooling and max-pooling are the two most widely used strategies, as illustrated by an example in Figure 2.3.

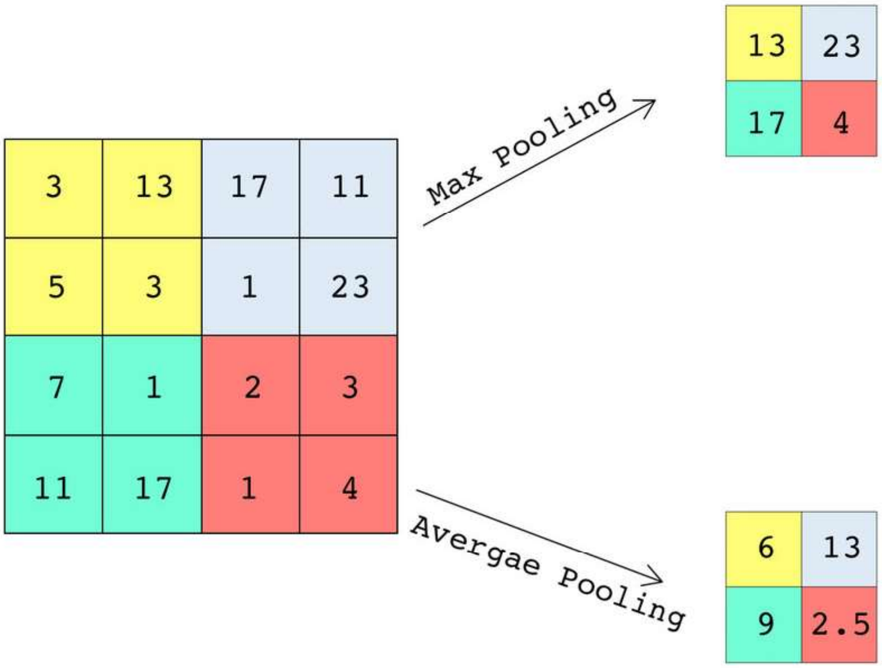

Figure 2.3: An example of a max-pooling and an average-pooling operation. In this example, a 4x4 image is down-sampled to a 2x2 image by calculating the maximum or average value of each sub-region.

Pooling layer is defined as:

$$\psi^2 = g(h(\psi^1)) \tag{2.2}$$

where $h(\cdot)$ is a pooling function and can be selected as max or average-pooling.

$\boxed{3.}$ Fully Connected Layers: They are used as the last few layers to better extract the information conveyed by lower layers in view of the final decision. We define fully connected layer ψ^{l+1} by considering the input feature map ψ^l, weights W^l and bias b^l for layer l as follows:

$$\psi^{l+1} = g(W^l \psi^l + b^l) \tag{2.3}$$

where $l = 2, 3, 4, \cdots$ and ψ^{l+1} denotes output of fully connected layer l.

2.1.3 Transfer learning

Transfer learning is a prominent method in computer vision, as it enables us to build precise models in a time-saving manner and is commonly used for predictive modelling problems that uses image data as input. With transfer learning, we basically try to exploit what has been learned in one task to improve generalization in another. It is difficult and time-consuming to collect very large medical imaging datasets of a specific domain. One of the problems associated with learning the CNNs parameter is that it requires a large amount of data and considerable computation power. When CNNs are trained on image data, such as ImageNet, we have the luxury of having a dataset of millions of labelled images that allow us to train relatively large networks with hundreds of layers to a high degree of precision. Thus, when the data is insufficient, one way to mitigate this is to apply transfer learning in supervised CNN models, pre-trained from a natural image dataset or a different medical domain [25]. A pre-trained CNN is applied to an input image in one scheme, and the outputs are then extracted from the network layers.

Training CNNs from scratch, however, needs a large number of images, or else the model will suffer from overfitting. A typical solution in these conditions is fine-tuning when only a part of the pre-trained neural network is fitted into a new dataset [26]. Fine-tuning is another approach used if a medium-sized dataset exists for the task. It uses a pre-trained CNN to initiate the network, and thereafter the training of several or all network layers is supervised using the new data for the task at hand [27]. Moreover, in many computer vision tasks, deep CNN architectures containing millions of parameters have achieved state-of-the-art results with fine-tuning [28, 29].

2.2 Related Works

In the mid-1980s, medical physicists and radiologists started to focus on computer-aided detection, and diagnosis [30]. Advances in computerized technologies and techniques such as machine learning and image analysis have played a critical role in improving diagnosis, sifting through areas that require treatment, and assisting professionals in their workflow throughout the last decade. The neural network has been extensively used in the development of research to explore intelligent ways to assist medical image diagnosis and has made significant progress. Deep learning, particularly CNN, has rapidly demonstrated critical skills for improving accuracy while also breaking new ground in data analysis, which is rapidly expanding. A CNN is a feedforward neural network that combines ReLU and pooling layers. It is typically made up of one or more convolutional layers as well as fully connected layers, as discussed in earlier section. In a classical CNN for image processing, convolutional filter layers are combined with pooling or data reduction layers. A convolution filter transforms a small portion of an image and can detect highly relevant visual features in a manner similar to the low-level pixel processing of the human brain. A CNN's output is typically the labelling of one or more probabilities or categories related to images. Convolutional filters can learn directly from trained data, which eliminates the need for time-consuming manual feature labelling. Many studies have shown that the neural network-based CAD technique outperforms the old method on tasks such as tumour segmentation and classification, implying that the neural network-based method has a broader scope and potentially higher value for clinical applications [31]. An overview of intelligent analysis of breast cancer histopathology has been presented in Table 2.1.

Table 2.1: Overview of the intelligent analysis of breast cancer histopathology.

Authors	**Year**	**Methodology**	**Conclusion**
Stenkvist et al. [32]	1978	Authors have proposed a new method for measuring differences in nuclear detail in chrome alum gallocyanin- stained nuclei of cells from HBCs and compared with conventional subjective grading and classification systems.	In general, imaging for cancer screening has been investigated for more than four decades. This study has demonstrated that individual cells obtained from human tumours by fine-needle aspiration biopsy are suitable for high-resolution image analysis.
Elston and Ellis [33]	1991	Proposed modified Bloom–Richardson system.	The histological tumour grade is commonly determined according to the modified Bloom–Richardson system, which consists of a semi-quantitative assessment of nuclear atypia, tubule formation, and mitotic activity.
Wolberg et al. [34]	1995	Authors estimated the classification of breast mass diagnosis based on fine-needle aspirates (FNA) by using digital image analysis and machine learning method.	Computer technology will improve breast fine-needle aspiration accuracy and prognostic estimations.

Continued on next page

Table 2.1 – *Overview of intelligent analysis of breast cancer histopathology*

Authors	**Year**	**Methodology**	**Conclusion**
Meijer et al. [35]	1997	Authors have given a brief overview of the origins of image analysis in pathology.	This review intends to give an overview of the developments that have led to image analysis as it is currently applied in routine diagnostic cytopathology and histopathology as well as in research.
Khoo et al. [36]	2005	Authors evaluated the recall and cancer detection rates prospectively with and without computer-aided detection (CAD) in the United Kingdom National Health Service Breast Screening Programme.	CAD increases the sensitivity of single reading by 1.3%, whereas double reading increases sensitivity by 8.2%.
Yang et al. [37]	2005	Authors presented a robust colour-based segmentation algorithm for histological structures that used image gradients estimated in the LUV colour space to deal with issues of stain variability.	They defined the colour gradient in LUV colour space in order to replace the grey level gradient of the original GVF snake. Although the capture range of the original GVF snake is large, it still may fail to find the edges of the object when given unsuitable initial positions.
Can et al. [38]	2008	Authors used mutual information-based error metrics to register the nuclei images from sequential staining steps.	A two-step technique was used to remove the AF from fluorescent microscopy images. Rather than acquiring images of all the dyes at once using a set of optimum filter cubes tuned to specific dyes, the acquisition is done in two steps.
Alomari et al. [39]	2009	Authors proposed a systematic approach to solve the localization problem in pathology images.	This method utilizes a feedforward backpropagation Neural Network with pre-and post-processing and many features, including; colour, texture, appearance, and location.

Continued on next page

Table 2.1 – *Overview of intelligent analysis of breast cancer histopathology*

Authors	**Year**	**Methodology**	**Conclusion**
Veta et al. [40]	2016	Authors developed an automated nuclei segmentation method that works with hematoxylin and eosin (H&E) stained breast cancer histopathology images, which represent regions of whole digital slides.	An accurate technique for automated segmentation of nuclei in images derived from digital slides of H&E stained breast cancer sections. The technique was evaluated on several representative regions and showed good performance in terms of detection and segmentation accuracy.
Irshad et al. [41]	2014	This survey covers major trends of various nuclei detection, segmentation, feature computation and classification techniques used in histopathology images, specifically in H& E staining.	It is shown that only a few supervised machine learning techniques are used for nuclei segmentation and feature extraction. So as future research work, more domain-specific feature extraction and efficient machine learning techniques can be explored.
Oquab et al. [42]	2014	Designed a method to reuse layers trained on the ImageNet dataset to compute mid-level image representation for images.	In this work, they showed how image representations learned with CNNs on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks with a limited amount of training data.
Spanhol et al. [14]	2016	Authors introduced a dataset composed of 7,909 breast histopathological images acquired on 82 patients. In the same study, the authors evaluated six different textural descriptors and different classifiers and reported a series of experiments with accuracy rates ranging from 80% to 85%, depending on the image magnification factor.	It is undeniable that the texture descriptors can offer a good representation to train classifiers. However, some researchers advocate that the main weakness of the current machine learning methods lies exactly in this feature engineering step.

Continued on next page

Table 2.1 – *Overview of intelligent analysis of breast cancer histopathology*

Authors	**Year**	**Methodology**	**Conclusion**
Spanhol and Caroline [43]	2016	Authors conducted some preliminary experiments using the deep learning approach to classify breast cancer histopathological images from BreaKHis, a publicly available dataset.	This method aims to allow use of high-resolution histopathological images from BreaKHis as input to existing CNN. The CNN performance is better when compared to previously reported results obtained by other machine learning models trained with handcrafted textural descriptors.
Vahadane et al. [44]	2016	Proposed a novel structure-preserving colour normalization (SPCN) scheme, which changes the colour of one image (source) to match that of another (target) while reliably keeping source structural information intact.	Color normalization is an essential step to remove undesirable colour variation in histopathological images and is useful for analyzing disease and its progression on large datasets from different pathology labs.
Bejnordi et al. [45]	2016	Presents and evaluates a fully automatic method for detection of ductal carcinoma in situ (DCIS) in digitized hematoxylin and eosin (H&E) stained histopathological slides of breast tissue.	The result of the experiments demonstrates the efficacy and accuracy of the proposed method as well as its potential for application in routine pathological diagnostics.

2.2.1 Feature learning

Feature learning consists of feature extraction, dimensionality reduction and learning models. Since our study focussed on the classification of cancer histopathology, thus we emphasized on classification techniques. These classification techniques are based on either machine learning, deep learning or a hybrid approach. Firstly we discuss basics and related works of feature extraction, dimensionality reduction techniques and thereafter, in the next section, classification approaches.

2.2.1.1 Feature extraction

For quantitative assessment of tissue and organ function extracting proper features is essential for characterizing cellular and tissue structures. In feature extraction, various features are unsheathed from the images representing their characteristic attributes. Owing to the differences in the characteristics of malignant and benign lesions, the extracted feature space is largely complex and massive. Methods for extracting features from histopathological images can be classified broadly into statistics based and learning-based methods. Different classical feature descriptors, such as color histograms [46, 47, 48, 49], Local Binary Pattern (LBP) [50], Scale-Invariant Feature Transform (SIFT) [51], and Histogram of Oriented Gradient (HOG) [52], have been explored for histopathological images. An operator often used for describing the textural features of an image is LBP, first developed by Ojala at Oulu University [50]. Further, it is reported in [53]. Completed Local Binary Pattern (CLBP), a different operator from the traditional operator, is proposed in [54]. This operator has 3 descriptors; CLBP- M, CLBP-S, CLBP-C. In the form of histograms, these descriptors can be fused in series, parallel or series-parallel, leading to a significant improvement in the textural features classification. In [55], histopathological image representation is given by the statistics relating to the size, shape, and distribution of the nuclei. Graph-based features [56, 57] were introduced for describing meaningful objects in the histopathological images like glands and ducts. Traditionally, different types of handcrafted textural features were extracted from breast images for the segmentation and classification of breast lesions. These texture features [58, 59] included size, shape, margin and intensity features. A summary of object-level features used for histopathological image analysis is presented in Table 2.2.

With advancements in deep learning methodologies, the entire process of features extraction is improved significantly. Mixed features [60, 61, 62, 63, 64, 65, 66] have been proposed for establishing more comprehensive representations, and showed superior per-

Table 2.2: Description of object-level features used in CAD of histopathology images.

Category	**Features**
Chromatin-Specific	Area, Integrated optical density, Mean optical density, Number of regions, Compactness, Distance, Center of mass
Radiometric and Densitometric	Hue
	Image Bands, Intensity
	Optical density, Integrated optical density, and Mean optical
Size and Shape	Area
	Boundary Features: Perimeter, Radii, Perimeter Fourier Energies, Perimeter curvature, Bending energy, Perimeter fractal dimension
	Bounding Box Features: Extent, Aspect ratio
	Center of Mass
	Convex Hull Features: Convex area, Convex deficiency, Solidity
	Elliptical Features: Major and minor axis length, Eccentricity, Orientation, Elliptical deviation
	Filled Image Features: Filled area, Euler number
	Other Shape Features: Equivalent diameter, Sphericity, Compactness, Inertia shape
	Reflection Symmetry
Texture	*Co-occurrence Matrix Features*: Inertia, energy, Entropy, Homogeneity, Maximum probability, Cluster shade, cluster
	Entropy
	Fractal Dimension
	Run-length Features: Short runs emphasis, Long runs emphasis, Gray-level non-uniformity, Run-length nonuniformity, Runs percentage, Low gray-level runs emphasis, High gray-level runs
	Wavelet Features: Energies of detail and low resolution images

formance. Kandemir et al. in [64], demonstrated mixing of a set of well-designed nuclear features with colour histogram, LBP, and SIFT. The necessities of patterns from colour, texture and nuclear statistics. Advancements in machine learning algorithms, various methods for feature learning were introduced like auto-encoders [67, 68, 69], sparse representation [70, 71, 72], restricted Boltzmann machines [73, 74], and CNN [75, 76, 77]. These methods were explored for mining the patterns from the raw histopathological image data. In [73], a restricted Boltzmann machine-based method was adopted for extracting sparse features from patches, and then a max-pooling process was utilized for quantifying through the image-level features. Sparse auto-encoders were applied in [67] for learning a set of feature extractors for histopathological images. Thereafter, convolutional operations on the feature extractors were utilized for obtaining feature maps. In addition, a stacked feature representation was added to the previous framework in [68]. In [76], Xu et al. designed a deep CNN for classification of the stromal and epithelial regions in histopathological images. Patch by patch extraction of features throughout the histopathological images ignored the characteristics of histopathological images and resulted in a high computational complexity in these approaches. Contrastingly a sparsity model is proposed in [71] for encoding cellular patches selected through a blob detection approach. Thereafter histopathological images were classified by fusing the predictions of these cellular patches. In continuation to this work, a class-specific dictionary learning method has been proposed in [72] for the sparsity model. This dictionary learning method was utilized for extracting more discriminative sparse representations of image patches. The patch-selected scheme used in [71, 72] sharply reduces the execution time, but many times fails in the identification of the malignant samples in cases minor malignant areas are present in the image.

In [78] authors concluded that deep features extracted from CNN trained over large datasets have a remarkable generalization ability for performing discrimination tasks and showed superior performance than traditional machine learning methods and methods utilizing hand-engineered features. In [79] authors reported a 7-layer deep CNN, and in [80] researchers proposed a 6-layer deep CNN for extracting features from mammograms. Many studies demonstrated that using pre-trained CNNs for features extraction leads to superior performance in comparison to conventional models based on handcrafted features.

Since the use of excessive features reduces the classifier performance due to redundancy, irrelevant as well as noisy features and also increases complexity, it is essential to filter the optimum feature set. Thus, learning of the most appropriate and crucial features is a prerequisite for the accurate image classification task [81]. In a way, feature extraction leads to dimensionality reduction by representing a compressed feature vectors. For measuring variations at the cell or tissue level, different types of features such as texture, morphological, co-localization and spatial features are utilized.

2.2.1.2 Dimensionality reduction techniques

The techniques for dimension reduction target to reduce dimensionality based on some other criterion than classification performance. Principal component analysis (PCA), linear discriminant analysis (LDA), and independent component analysis (ICA) are the three most common techniques for reducing the dimensions of input data feature matrix. PCA [82] tries to adopt a new orthogonal coordinate system wherein the maximum data variance is included in the first few dimensions. Varying degrees of variance is observed upon the projection of the data. The maximum variance in the data is encompassed by the first coordinate, thereafter by the second coordinate and so on. Unlike PCA, LDA is a supervised method requiring class labels for individual data samples. It thus maps the data onto a lower-dimensional sub-space for better data discrimination. The main goal is to find the map wherein the sum of distances between samples in the same classes is minimized while maximizing the sum of distances between samples in different classes. In ICA [82], the aim is to formulate a mixing matrix so that the mixture of the features are independent statistically. Performance of ICA for extracting three independent components, i.e., three tissue types, has been demonstrated for segmentation of colon histology hyperspectral images [83]. Unlike PCA, ICA doesn't give a ranking of the independent components. A number of other methods are also available for calculating the independent components, which are, however, computationally intensive.

2.2.2 AI approaches for histopathological image classification

Most of the studies conducted on histopathology breast cancer classification have utilized deep learning, machine learning or a hybrid of both AI techniques to achieve significant results. One of the interesting facts in histopathology images that makes it different from other medical images is that available breast cancer histopathology images are collected on different magnifications. Therefore we divide AI approaches used to analyze breast cancer histopathology into magnification specific classification approaches that include both machine learning and deep learning approaches as well as magnification independent approaches.

2.2.2.1 AI approaches for magnification-dependent model

It has been observed from various literature that 85% of CAD systems for BreakHis followed the magnification specific formulation for tumour classification.

1. **Machine learning approaches based on handcrafted features:**
   CAD systems first developed for the BreakHis dataset were based upon a dual-stage traditional approach, whereby handcrafted features were extracted from the images and were used for training a standalone classifier. In [14] the researchers studied the efficacy of 6 state-of-the-art handcrafted features descriptors namely: LBP [50], its variant CLBP [53], Gray Level CoOccurrence Matrices (GLCM) [84], Local Phase Quantization (LPQ) [85], Oriented FAST and Rotated BRIEF (ORB) [86] and Parameter-Free Threshold Adjacency Statistics [87]. The performance

of these feature descriptors was evaluated along with four different classifiers: Support Vector Machines (SVM) [88], 1-Nearest Neighbor (1-NN) [89], Quadratic Linear Analysis (QDA) [90], and Random Forests [91]. In [92], researchers trained an SVM classifier with the fractal dimension [93] of each image and proved that using fractal dimension as a unique feature descriptor is more appropriate for classifying lower magnification (40x) images, which are having a number of self-similarities; however it is of little use for higher magnification images having less self-similarities.

Authors in [92] also attempted multi-category classification using 16 experiments, each classifying malignant and benign sub-classes. In [94], authors experimented with various handcrafted descriptors in combination with a k-NN classifier, including GLCM, LBP, PWT and TWT. In [95], an L1-norm sparse SVM (SSVM) [96] is proposed for selection of the most relevant features from the histopathological images in the BreakHis dataset. They observed that, L1-norm is unreliable for precise feature selection, and the SSVM showed biasness towards large hyperplane coefficients. For improving the quality of feature selection, a weight was assigned to each feature based upon its Wilcoxon rank-sum [97]. In [98], authors tested performance of KAZE features [99] using a bag-of-features approach.

Limitations: The traditional handcrafted features demonstrated relatively acceptable but highly unstable preliminary results. The main concern with these traditional approaches is that the model's quality is dependent on the feature extraction task, and obtaining representative features in a real sense is a very difficult task. A major difficulty is in choosing the appropriate descriptor. Even after combining different descriptors for enhancing their discriminative power, or their post-transformation for selecting the best ones, the results remain somewhat inferior and are sometimes not stable between different magnification levels.

2. **Deep learning-based approaches:**

Researchers in [6] were the first to evaluate a deep learning-based system for this BreakHis dataset by a CNN model incorporating features extraction and classification tasks. First, they explored LeNet [100], and found that it gives poor performance as compared to traditional methods in [14]. Then, researchers tried AlexNet [101] as a comparatively deeper network. Though CNN models use raw images as input, but in [102] authors utilized texture and pixel distributions in handcrafted features obtained using LBP or histogram descriptors. Upon comparison of different combinations, they achieved the best results with the Model 1-CNN-CH model. This model is based upon CNN architecture with residual blocks derived from ResNet [103], and involves Contourlet Transform (CT), and Histogram information descriptors for concatenation of extracted local-features [104]. Different popular schemes in deep learning utilized for histopathology breast cancer analysis are detailed as below.

(i) **Autoencoder:** A method of unsupervised learning, "Back-propagation", is used by the autoencoder, which adjusts the target values to match the inputs. It's a three-layer neural network containing the input layer, hidden layer, and decoding layer. An autoencoder is a device for converting input data into a hidden layer, which is then reconstructed by the decoder. Denoising autoencoders (DAE), sparse autoencoders (SAE), variational autoencoders (VAE), and contractive autoencoders (CAE) are the four types of autoencoders. By lowering the data's dimensionality, auto-encoders provide us with an edge. They can do both linear and non-linear transformations. The training of an autoencoder requires considerable volumes of data, execution time, hyperparameter tweaking, and model validation. For classification of nucleus patches on HBC histopathological images, Xu et al. [105] presented a stacked sparse autoencoder (SSAE) framework consisting of two SAE. For nuclei patch classification, SSAE+ Softmax performed better than the traditional softmax classifiers, PCA+Softmax, and SAE+Softmax. The breast cancer images, though, in this study were taken from only 17 patients. Following that, authors in [69] suggested a method for detecting breast cancer nuclei using the SSAE framework on 537 H&E stained histological images. However, they did not examine any pre-processing procedures. Kadam et al. [106] suggested feature ensemble learning with SSAE and showed that it outperformed the SSAE+softmax framework.

(ii) **Transfer learning:** It is a popular method in computer vision since it allows us to quickly develop exact models. It's widely used for predictive modelling issues that employ image data as input. The aim of using transfer learning is to use what we have learned in one task to boost generalization in another. Collecting very big medical imaging datasets of a specific domain is difficult and time-consuming, and in such cases, transfer learning is of great help. CNNs require a lot of data and processing power. Training of CNNs on image data, such as ImageNet, which is a database of millions of annotated images, allows training rather large networks with hundreds of layers with excellent precision. In cases where data is insufficient, one solution is to use transfer learning in supervised CNNs models that have been pre-trained on a natural image dataset or in a separate medical domain [25]. In one technique, an input image is fed into pre-trained CNNs, and the outputs are recovered from the network layers. If a medium-sized dataset is available for the task, fine-tuning is another option. It starts the network with pre-trained CNNs and then supervises the training of some or all network layers using new data for the task at hand [27]. Furthermore, fine-tuning deep CNNs architectures with millions of parameters have obtained state-of-the-art performance in numerous computer vision problems [29, 28]. However, a vast number of images are required for training CNNs from scratch to prevent the model from overfitting.

(iii) **CNN-based approaches:** CNNs were used for the 2015 Bioimaging breast histology classification challenge dataset for binary and multi-class (normal/benign/in situ/invasive) classification of images with an accuracy of 83.3% and 77.8% respectively by applying rotation and mirroring as data augmentation strategies for both tasks [107]. Pre-trained CNN can be used for feature extraction as well as for fine-tuning. Different approaches are utilized for fine-tuning of pre-trained CNN. Some approaches proposes fine-tuning of all the layers in pre-trained CNNs, whereas, in some, only the last fully connected layer is used for pretraining. A strategy based upon dual-stage fine-tuning is adopted in [108] proposed, wherein in the first stage, only the fully connected layers are retrained, followed by training of the whole network. The authors also evaluated the two stages independently and found that the dual-stage model performs better.

Features extracted from a pre-trained CNN using ImageNet are supposed to be the high-level features between ImageNet and BreakHis classification tasks. But for BreakHis classification, the used CNN is not supervised for extracting features, generating a gap between the extracted features and the needed domain-specific features [109]. For reducing this gap, authors in [110], proposed a deep domain knowledge-based feature model, which is a hybrid transfer learning approach.

In [111] authors explored an approach based upon transfer learning using a pretrained AlexNet and DeCAF based features extraction [78]. The features were extracted from the last layers of pre-trained AlexNet's and were then used for training a standalone classifier. Thereafter in [112] authors evaluated three different dimensionality reduction methods i.e., PCA [83], Gaussian Random Projection (GPR) [113] and Correlation-Based Feature Selection (CBFS) [114] on the features extracted from a pretrained VGG [115]. In [116] authors introduced a sequential features extraction framework in combination with XGBoost classifier [117], for evaluating the potential contained in each layer of a pre-trained DenseNet169 [118]. The results showed that compared to the final fully connected layers, the last convolutional layers provide more significant features. Interestingly it was also observed that the lower-level layers contribute significantly for classification of low magnification (40×) images and mid-level features are useful for 100× and 200× magnifications, whereas higher magnification images (400x) are effectively represented by higher-level layers.

For the 2015 Breast Cancer Classification Challenge, the authors in [119] developed a methodology for classification tasks using Inception version 4, residual network transfer learning frameworks, and a recurrent CNN architecture. They artificially increased the size of the datasets using data augmentation methods. Experimental results demonstrated an accuracy of 97:57 ± 0:89% for multi-class and 97:95 ± 1:07% for binary classification of cancers using the BreakHis dataset. To predict malignancy and image magnification levels, Bayramoglu et al. [120] used single and multi-task CNN architectures, and to supplement the dataset, cropping and rotation were used on the BreakHis dataset. They attained an accuracy of 82.13% utilizing multi-task CNN architecture and a classification rate of 83.72% using a single job for the benign/malignant binary classification task.

For binary breast cancer classification, Kassani et al. [121] used an ensemble of models based on transfer learning (VGG19, MobileNet, and DenseNet) on 4 benchmark datasets, namely BreakHis, PatchCamelyon, 2015 Bioimaging Challenge, and 2018 ICIAR datasets with accuracies of 98.13%, 94.64%, 95%, and 83.1% respectively. Data augmentation techniques such as flipping, zooming, shear, and rotation were also applied.

On the BreakHis dataset, authors in [122] employed a deep learning model integrating bi-LSTM and fully connected network architectures for binary classification and achieved an accuracy of 96.32%. In [123] researchers utilized a deep learning-based model to classify histopathological images into benign/malignant categories. They used belief theory to conduct experiments on the BreakHis dataset, merging ResNet-18, ResNet-50, and AlexNet architectures, obtaining image-level accuracy of 96.88%. Sudharshan et al. [10] used BreakHis dataset histopathology images to develop a weakly supervised method for binary categorization (benign/malignant) of tumours with an accuracy of 92.1% at 40x magnification. The method works well without the need for labelled images, which is a significant benefit of their method. On the BreakHis dataset, Spanhol et al. [111] suggested a deep learning model that utilized a previously trained CNN model, attaining a subject level F1-score of 90.3%. In [124], a matrix power normalization approach is presented by authors to include global covariance information into a simple CNN model. By utilizing second-order statistical information, the framework resulted in effective representations from histology images on the BreakHis dataset with 97.92% accuracy at the subject level for the binary classification task. Data enhancement techniques such as cropping and flipping were also utilized by the authors to artificially increase the size of the dataset. In [123], authors employed BreakHis and a custom dataset to classify benign and malignant tumour cells using several transfer learning architectures (GoogLeNet, VGGNet, and ResNet). They used data augmentation methods such as scaling, translation, colour augmentation, and rotation, to attain a classification accuracy of 97.67%.

- **CNN architecture adaptation**

  Authors in [125], presented "Transition module", a newer version of GoogleNet inception module, and its integration with AlexNet. The updated version was created to make the transition from the last convolution layer to the first completely connected layer less abrupt. This transition module, unlike the inception module, does not incorporate any prior dimensionality reduction. A new CNN architecture made up of fifty convolutions is presented and compared to numerous handcrafted feature-based models [126]. BiCNN, a new CNN based on the GoogleNet architecture, is proposed in [127] and sub-class information along with binary labels of each image is utilized for using it in binary classification task. Taking into account both annotation levels would help the proposed model to learn features distance more effectively between binary classes.

- **Deep features extraction versus traditional features extraction**

  Authors in [128] compared features extracted with traditional handcrafted descriptors and those obtained from a pre-trained AlexNet. The authors found that the performance of LBP handcrafted feature extraction is somewhat better than the AlexNet features. The comparison, however, was very restrictive with a somewhat shallow CNN, which in [6] was also found to hardly outperform handcrafted based models.

Limitations: The results obtained using various deep learning models are far better in comparison to standard methods. Deep learning models, on the other hand, are particularly data-hungry and demand a big amount of data, whereas medical applications like breast cancer diagnosis are usually short on data. To get rid of this, researchers are frequently required to use artificial data augmenters as a pre-processing step. Furthermore, there is no guiding principle for choosing the appropriate hyper-parameters for these types of models. Moreover, unlike manually constructed models, deep learning algorithms are unable to provide users with feedback on the discriminative features that are utilized to determine each individual's diagnosis. Aside from HBC, there are several deep learning opportunities and a variety of web-based medical applications, each with its own set of constraints [129].

$\boxed{3.}$ **AI approaches for magnification-independent model:**

The authors offered a magnification-independent strategy for the BreakHis classification task in [120], wherein they advocated classifying histopathology images as benign or malignant without magnification variables. They devised two experiments to assess their magnification independent approach and investigated a single-task CNN in the first and trained a multi-task CNN in the second. For comparing with earlier magnification-dependent research, the single task CNN was trained on all pooled magnification subsets and then tested on every magnification separately. The study revealed that the magnification-independent model outperforms several earlier magnification-dependent research and produced consistent results across magnifications. The authors investigated the performance of a multi-task variant of the first magnification-independent CNN in their second experiment. An additional classifier was added to this multi-task version to estimate the magnification level of each input image. When compared to the results achieved with the single job version, its binary classification results declined marginally. The employment of the extra magnification factor classifier for the binary classification problem justified the loss in accuracy.

With a cross-magnification training/testing strategy, the authors of [130] attempted to examine the discriminative utility of each magnification subgroup individually. The authors attempted to evaluate each feasible training/test combination, whereby the model is given training on a subset, and then testing is done on the same subset or a different subset. Each model used a two-stage approach. It began by extracting colour and texture attributes, and thereafter it combined these features to create a majority voting ensemble that included SVM, Nearest Neighbors, Discriminant Analysis, and Decision Tree.

4. **Magnification-dependent vs Magnification-independent model:**
In practice, we could either take the magnification factor of each image into account while training a magnification-dependent classification model or train it regardless of the magnification feature. Training model with the magnification-independent approach is more suitable than using a magnification-dependent strategy due number of factors.

    (i) With a magnification-dependent approach, one model is required for each magnification subset, requiring more training and adaption.

    (ii) When using a magnification-dependent technique to test samples, the magnification factor of each test image must be specified before applying the associated model. This specific information, however, may not be available for all images.

    Unlike a magnification-dependent model, a magnification-independent model can directly benefit from extra training data, which can be obtained using the same or alternative magnification factors. The authors of several magnification-dependent models [131, 132] got better results when they used ensembles with 4 magnification-dependent models. This fact might be viewed as support for the concept that training a model with features from many scales (in a magnification-independent manner) can help it to generalize better. The magnification specific approach was embraced by the majority of works since it is the most straightforward; however, to date, magnification-independent is the greatest clinical and practical reformulation for BreakHis. In fact, because of the difficulty of learning characteristics from all magnification levels at the same time, this reformulation has never been used in any previous work. Furthermore, categorizing each image into one of eight separate classes with comparable patterns is a time-consuming process.

5. **Private and secure classification Model for cloud-based system :** Nevertheless, the advent of the cloud and the collaborative model raises privacy concerns. Privacy concerns are related specifically to confidential input data during training or inference, as well as to the sharing of a trained model with other people. Security threats to AI systems may include influence attacks, security violations, attack specificity, and adversarial attacks [133, 134]. Influence attacks attempt to gain control over the training data, whereas security violation may be an attack on integrity wherein the model is provided with harmful inputs or a privacy violation attack involving the disclosure of sensitive and confidential training data. In adversarial attacks, the adversary generates small, carefully crafted (unnoticeable) perturbations to the actual (non-modified) input samples to attack the integrity of the deep learning system. Adversarial attacks have emerged as a major security threat to machine learning/deep learning systems [135, 134, 136, 137, 138]. The discovery that deep learning models are not secure hinders their practical use in safety-critical applications such as predictive healthcare. Thus, in the modern era of the health IoT and Cloud-based clinical diagnostic model, security and privacy of the data and of the prediction model without compromising its accuracy has emerged as an open research area. Recent research also validates this point by proposing private machine learning or privacy-preserving deep learning models [139]. However, the collection of data for deep learning raises serious privacy concerns. Training data is usually sensitive and may originate from multiple sources that have different privacy requirements. A cooperative learning framework has also been reported by some researchers whereby data owners share a secure version of their data with each other and collectively train a global model over these data. Such studies on privacy-preserving machine

learning examine many models like Naive Bayes classification [140], k-means clustering [141], random decision trees [142, 143], and neural networks [144]. But these solutions appear to involve either high computational or communication overheads for data owners. Secure multiparty computation (SMPC) is a well-accepted way to perform operations in untrusted settings without the disclosure of data. In machine learning models, SMPC secures model weights while allowing multiple work points to contribute in the training phase with their own datasets, a method known as Federated Learning (FL). However, securely trained models tend to be vulnerable to reverse-engineering attacks that can specifically retrieve sensitive dataset information from the model. This issue is resolved by Differentially Private (DP) methods, which can protect data effectively. With the help of SMPC, even mutually distrustful parties can compute publicly known algorithms without disclosing their inputs. MPC has rapidly evolved from being a theoretical concept three decades ago [145, 146] to reality for solving real-world problems. In multiple instances, MPC has been shown to safely draw inference from machine learning algorithms where model and query are hidden from the participants of the protocol. There are many recent approaches for running inference securely with MPC like Gazelle [147], HyCC [148], Secure machine learning [149], MinioNN [150], CHET [151], Secure NN [152], ABY3 [153], and so on. However, these approaches are difficult to use for machine learning developers and have only been demonstrated on small deep neural networks using tiny datasets.

2.3 Summary

Based on literature review it can be concluded that though a number of studies have been conducted on CAD of HBC, but none of them is entirely convincing and suffers from shortcomings such as lack of extensive analysis regarding effect of stain normalization, illumination, magnification and other variables on the performance of AI models.

Since AI is a rapidly evolving field, latest developments in the area have opened up new avenues for developing improved models with higher accuracy then state-of the-art. Further, deep learning models are resource and data-dependent, and have poor applicability on resource constrained devices, while most of the clinics don't have such high compute resources. However very few studies have focussed on developing a framework that works efficiently on low-cost clinical settings with minimal data, and on mobile edge devices. Data privacy and model security concerning development of CAD model from clinical data are also not explored. CMT is an important neoplastic disease of female dogs and excellent model for HBC as well, however no studies have been conducted in the area of CAD of dog cancers due to lack of any publicly available dataset.