[Iraqi Journal for Computer Science and Mathematics](https://ijcsm.researchcommons.org/ijcsm)

Volume 6 | Issue 1 Article 1

2025

A Review of Breast Cancer Histological Image Classification: Challenges and Limitations

Israa Faisal Jassam
Registration and Students Affairs, University Headquarter, University of Anbar, 31001, Ramadi, Anbar, Iraq

Abdulrahman Abbas Mukhlif
Registration and Students Affairs, University Headquarter, University of Anbar, 31001, Ramadi, Anbar, Iraq

Ahmed Adil Nafea
Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi, Iraq, ahmed.a.n@uoanbar.edu.iq

Mustafa Adnan Tharthar
Registration and Students Affairs, University Headquarter, University of Anbar, 31001, Ramadi, Anbar, Iraq

Ahmed Isam Khudhair
Office of the Assistant President for Administrative Affairs, University of Anbar, Ramadi, Iraq

Recommended Citation

Jassam, Israa Faisal; Mukhlif, Abdulrahman Abbas; Nafea, Ahmed Adil; Tharthar, Mustafa Adnan; and Khudhair, Ahmed Isam (2025) "A Review of Breast Cancer Histological Image Classification: Challenges and Limitations," *Iraqi Journal for Computer Science and Mathematics*: Vol. 6: Iss. 1, Article 1.
DOI: https://doi.org/10.52866/2788-7421.1232
Available at: https://ijcsm.researchcommons.org/ijcsm/vol6/iss1/1

IRAQI JOURNAL FOR COMPUTER SCIENCE AND MATHEMATICS 2025;6:1–23

REVIEW

A Review of Breast Cancer Histological Image Classification: Challenges and Limitations

Israa Faisal Jassam [a], Abdulrahman Abbas Mukhlif [a], Ahmed Adil Nafea [b,*], Mustafa Adnan Tharthar [a], Ahmed Isam Khudhair [c]

[a] Registration and Students Affairs, University Headquarter, University of Anbar, 31001, Ramadi, Anbar, Iraq
[b] Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi, Iraq
[c] Office of the Assistant President for Administrative Affairs, University of Anbar, Ramadi, Iraq

ABSTRACT

This paper comprehensively reviews the classification of breast cancer histological images. The paper discusses the research objectives, methodologies used, and conclusions drawn, as well as suggestions for the future. The study is based on the ICIAR 2018 database, which is considered one of the largest databases available to support this research. The paper also addresses major challenges such as lack of data, variation in tissue preparation, class imbalance, and computational requirements. Advanced techniques such as deep learning (DL), transfer learning and data augmentation are explored, along with innovative models such as convolutional neural networks (CNNs) and generative adversarial networks (GANs). The study focuses on integrating multimodal (histological, genetic, and clinical) data to enhance diagnostic accuracy and enable personalized treatments. Breast cancer is responsible for 25.4% of cancer cases, with 1 in every 12 women developing this fatal disease. In 2020, a large number of women died as a result of the disease spreading through lymphatic and blood vessels, with the number of deaths reaching about 685,000 cases. Improving outcomes depends greatly on early detection of the disease and providing appropriate and effective treatment. Future directions include addressing challenges related to texture image classification and promoting standardized data sets with the goal of developing diagnostic tools and treatment strategies.

Keywords: Breast cancer, Deep learning, Machine learning, Classification, ICIAR 2018, Detection, Convolutional neural network

1. Introduction

This study seeks to explore methods for classifying tissue images of breast cancer, and will address the challenges and advances that researchers and medical personnel face in accurately diagnosing and treating breast cancer. Breast cancer is one of the most common types of cancer among women, and its diagnosis and treatment is done using histological images [1–4].

From the statistical chart in Fig. 1 of cancers in the world, we notice that breast cancer occupies the largest position, as breast cancer in women has surpassed lung cancer as the most commonly diagnosed cancer, with the number of new cases estimated at about 2.3 million cases (11.7%) [5].

On the other hand, identifying breast cancer is a complex task and requires great expertise in the field of medical classification [6]. To learn how to classify breast cancer in the correct ways, we will explain it in the following paragraphs:

Determining the stage of cancer is an essential element of pathology when studying the spread of breast cancer and the various treatment options [7]. Digital imaging techniques have evolved and histological images have become an important means for diagnosing and staging breast cancer [8] On the other hand, some obstacles and limitations must be overcome to ensure

Received 29 September 2024; revised 7 January 2025; accepted 10 January 2025.
Available online 5 February 2025

* Corresponding author.
E-mail address: ahmed.a.n@uoanbar.edu.iq (A. A. Nafea).

https://doi.org/10.52866/2788-7421.1232

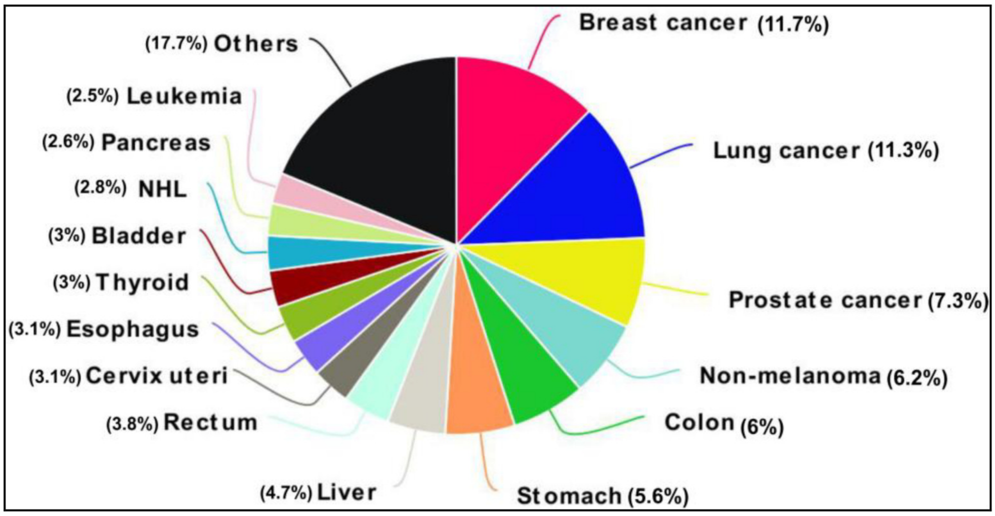

Fig. 1. Global cancer statistics (2020), showcasing breast cancer's prominence among women (Data from GLOBOCAN 2020) [5].

accurate and trustworthy classification for breast cancer detection [9].

Manual classification of breast cancer on histopathological images remains difficult due to some limitations. The examination requires the professional background and rich experience of pathologists, which may increase the cost of diagnosis time or cause error. Then, automated computer-aided diagnosis (CAD) in analyzing medical histology images of breast cancer is essential and greatly important to guide treatment and avoid medical errors for pathologists [10]. Many studies have shown that the use of histological images in diagnosing and staging breast cancer is of great benefit [11]. Doctors and researchers are working to improve breast cancer staging using tissue images to guide treatment more effectively and accurately [12].

The main objectives of this study are:

- Highlighting the most important models used to classify breast cancer tissue images.
- We aim to highlight the complexities that researchers and practitioners face in accurately evaluating and staging breast cancer through histopathological analysis.
- By accurately classifying these images through histopathological analysis, healthcare professionals can make more informed decisions regarding patient care, treatment planning, and prognosis.

2. Accurate evaluation of breast cancer staging using histopathological analysis

To identify different breast cancer subtypes, histological images obtained from patients should be properly analyzed and evaluated [13]. After following this procedure, the diagnosis of the disease can be determined more accurately, treatment options can be chosen and the general condition of the patient can be managed more accurately [14]. The goal of this procedure is to classify cancer cells more accurately according to their characteristics (morphology, size, and coloring patterns). Thus, this classification helps health care professionals in making the appropriate decision about diagnosis and providing appropriate treatment [15].

Histological images for classifying breast cancer are very important to show the cellular structure of breast tissue, which in turn is an important step for classifying breast cancer to choose the appropriate treatment for patients [16]. Thus, medical experts can distinguish the various features of breast cancer cells' analyzing these images, including Features (shape, dimensions and organization). These features allow them to evaluate the severity of the disease and divide breast cancers into different subcategories.

Fig. 2 shows the microscopic patterns of benign breast tumors. Figs. 3 and 4 represents the microscopic patterns of malignant breast tumor [17].

Histological images can be visually examined to find patterns and abnormalities that can be classified. This is all made easy thanks to the use of machine learning (ML) algorithms and computer-aided design (CAD) programs. Diagnosis can be improved using these modern tools, which can analyze huge amounts of images very quickly and with greater accuracy than human methods [18–20].

In order to allow for improved diagnosis and personalized medical treatment, this can be done by combining both histological images and genetic and clinical information, as it provides a more

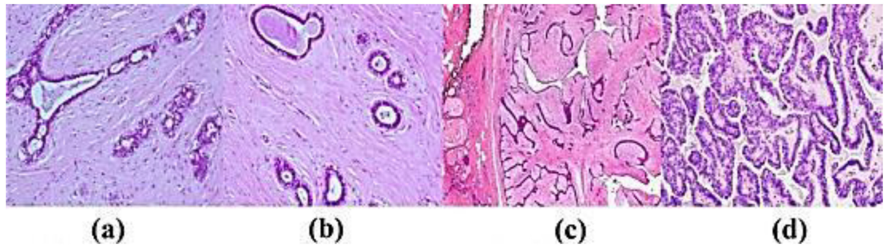

Fig. 2. Microscopic patterns of benign breast tumor (a) Fibroadenoma (Intracanalicular pattern), (b) Fibroadenoma (Pericanalicular pattern) (c) Phyllodes tumor (d) Intraductal papilloma [17].

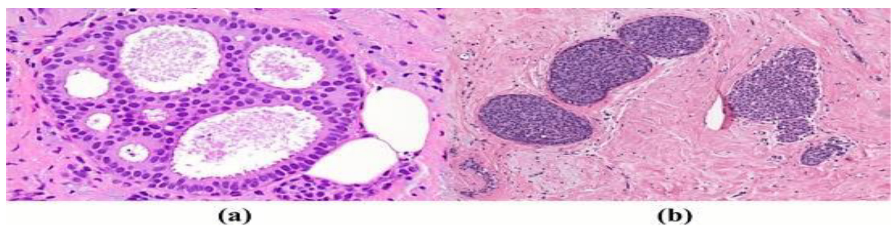

Fig. 3. Microscopic patterns of Noninvasive (In situ) carcinoma (a) Intraductal carcinoma (b) Lobular carcinoma [17].

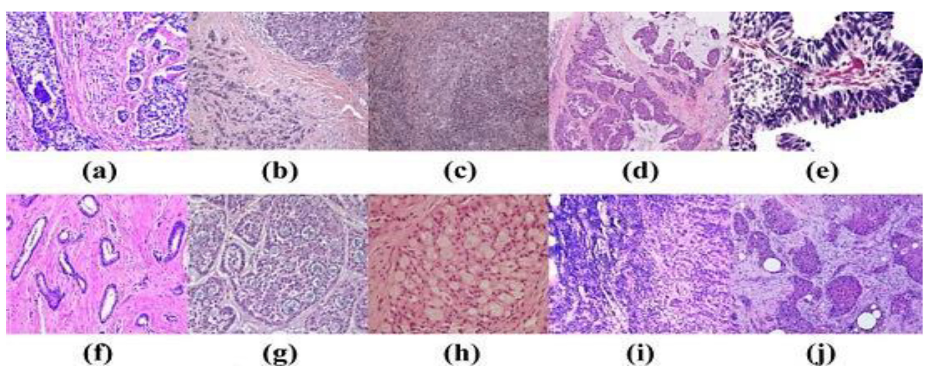

Fig. 4. Microscopic patterns of Invasive carcinoma. (a) IDC (b) Invasive lobular carcinoma (c) Medullary carcinoma (d) Mucinous carcinoma (e) Papillary carcinoma (f) Tubular carcinoma (g) Adenoid cystic carcinoma (h) Secretory carcinoma (i) Inflammatory carcinoma (j) Carcinoma with metaplasia [17].

comprehensive understanding and integration of breast cancer. Recent evidence that supports this integration of research includes [21–23]:-

1. Improve diagnostic accuracy: To obtain high accuracy for breast cancer subtypes, this is done by revealing underlying tumor characteristics by integrating genetic markers with histologic patterns and the patient's clinical history.
2. Integration of genomic data enhances diagnosis and treatment planning (such as chromatin accessibility and mutation profiles) by complementing histologic analysis to improve diagnosis for targeted therapies.
3. To create robust predictive models by combining multimodal data integration with genomic and clinical records of breast cancer histopathology images.

The result of improving the classification of estrogen receptor-positive breast cancers came as a result of combining histological images and chromatin accessibility data, as shown by Xu et al. [21]. For robust classification, Xiao and Lu [22] in their paper presented a framework that integrates labeled and unlabeled datasets. To prove that integration helps in model adaptation and classification performance, Mukhlif et al. [23] highlight the potential of dual transfer learning for merging unlabeled and labeled medical images.

2.1. Breast cancer classification overview

Breast cancer affects millions of women globally, which can be complex or diverse. When cancer cells in the breast multiply abnormally, they produce tumors. The patient's survival rate and the treatments given to him are also affected by the early diagnosis and classification of the disease [24–26].

2.2. Types of breast cancer

Invasive lobular carcinoma and ductal carcinoma, as well as ductal carcinoma in situ (which is the most common) are types of breast cancer. The type of cancer affects diagnosis and treatment [27].

2.3. Methods of classification

Breast cancer can be classified using ML, deep learning (DL), and feature extraction techniques, which have their advantages and disadvantages [28]. Using pre-trained models such as ResNet and Xception, which leverage transfer learning, this in turn plays a critical role in breast cancer image classification, thus reducing the requirement for large labeled datasets. Researchers such as Mukhlif et al. [23] are improving methodologies by fine-tuning models on labeled and unlabeled data.

Xiao and Lu [22] show that combining labeled and unlabeled data improves classification by recognizing complex patterns. By using techniques such as clustering, unsupervised learning helps take advantage of unlabeled medical images.

S. Xu et al. [21] revealed that multimodal data integration between histopathological images, genetic and clinical data enhances diagnostic accuracy and deepens understanding of disease mechanisms.

2.3.1. Classification of histological images of breast cancer patients using ML algorithms

Classification of breast cancer histological images is the cornerstone of research and progress in the field of medical imaging. Machine learning techniques are essential for accurate and efficient classification. To obtain accurate and effective classification, well-known ML algorithms must be used.

These algorithms include: -

- KNearest Neighbors (K-NN),
- Random Forest (RF),
- Naive Bayes (NB), and
- Support Vector Machine (SVM).

The above algorithms use information extracted from histological images to classify them as cancerous or benign. The characteristics that are extracted from these images are: the shape, texture and density of the images.

SVMs are widely used in breast cancer classification due to their ability to handle high-dimensional data and complex classification problems. Integrating several decision tree projections using RF has yielded impressive results in this field. While K-NNs use NBs to classify unknown data, NBs use probabilistic relationship between features and class labels [29].

2.3.2. Diagnosis of histological images of breast cancer patients using deep learning (DL)

Deep learning (DL) has shown significant improvements in breast cancer diagnosis over traditional methods such as classical machine learning (ML) and manual examination, and DL is considered a transformative technique. Early ML approaches generally relied on algorithms such as random forests (RF) and support vector machines (SVM), which require handcrafted features from histopathological images. Although these methods are effective, they lack the ability to scale or adapt when the dataset is complex and diverse [30, 31].

The introduction of convolutional neural networks (CNNs) has been a major advance in improving feature extraction and enabling end-to-end learning. Models such as AlexNet and VGGNet have demonstrated their ability to achieve high classification accuracy by taking advantage of large datasets such as ImageNet [32]. Reinforced transfer learning optimized these models by modifying structures pretrained on specialized breast cancer datasets, thus reducing the requirement for large amounts of labeled data [33].

Recent innovations in DL for breast cancer diagnosis include:

1. **Hybrid Architectures:** Merging CNNs with Recurrent Neural Networks (RNNs) to preserve both spatial and temporal characteristics of histopathological images [34].
2. **Generative Adversarial Networks (GANs):** through creation of synthetic images can tackling data shortage and class imbalance [35].
3. **Attention mechanisms:** To improve interpretation ability and diagnostic accuracy, critical areas in the images must be highlighted [36].

2.3.3. Diagnosis of histological images of breast cancer patients using transfer learning

Breast cancer diagnosis using transfer learning has made significant progress, using pre-trained models to effectively classify histopathological images. Following are some recent studies:

1. **Dual transfer learning approach:** Mukhlef et al. [23] proposed a dual transfer learning approach that involves training models on unlabeled medical images before optimizing them on labeled datasets. This has enhanced the flexibility and classification accuracy of breast cancer tissue images.

2. **Combining supervised and unsupervised learning:** Xiao and Luo [22] successfully propagated labels from a small subset to unlabeled images, by integrating unsupervised clustering into semi-supervised frameworks, as well as addressing and tackling variability and imbalance in extensive datasets.
3. **GAN-enhanced Transfer Learning:** Nassima et al. [35] explored the use of Generative Adversarial Networks (GANs) to augment limited datasets. Their intra-domain fine-tuning method bridged gaps between histological datasets, boosting classification precision.
4. **Transfer learning with hybrid architectures:** Yan et al. [37] introduced a hybrid model that integrates convolutional and recurrent neural networks, and effectively absorbs and captures local and global image features to improve breast cancer classification.
5. **Lightweight CNN Models**: Kausar et al. [36] suggested a streamlined deep CNN model tailored for environments with limited database, emphasizing efficiency and lower computational expenses while maintaining accuracy.
6. **Integration of multimodal data:** To improve diagnosis and provide a comprehensive view of estrogen receptor-positive breast cancers, Xu et al. [21] proposed integrating histological imaging with chromatin accessibility data.

These studies demonstrate innovative uses of transfer learning to improve diagnostic accuracy, reduce reliance on large datasets, and effectively manage data heterogeneity.

2.3.4. Diagnosis of histological images of breast cancer patients using unsupervised learning techniques

Unsupervised learning techniques have increasingly been applied to breast cancer tissue imaging diagnosis. Some recent studies are as follows:

To augment synthetic data, Farid-Adar et al. [38] used GANs and unsupervised support learning indirectly in medical imaging.

To extract significant patterns from histopathology images, Xingyu Li et al. [39] developed a technique using a fully convolved autoencoder, which supports breast cancer classification without requiring extensive annotations.

To aid in the automated detection process of breast cancer diagnosis, Bilal Ahmed Lodi [40] proposed an unsupervised method that uses hierarchical clustering to locate masses in mammograms.

To demonstrate the effectiveness of automated clustering in identifying cancerous regions, Sangwon Lee et al. (2020) presented an unsupervised learning-based approach for detecting for cancer detection in invasive breast cancer slide images [41].

Alexander Thiari et al. [42] demonstrated a colorless algorithm that learns independently from pathology images, this algorithm improves the robustness of the algorithm in dealing with color variations and increases the performance of the algorithm on breast cancer datasets.

To extract features and discover patterns, Zhao et al. [43] examined the use of deep learning in histopathology, highlighting unsupervised approaches.

In order to improve the efficiency of texture image classification, Xiao and Lu [22] developed a semi-supervised framework that integrates unsupervised clustering.

These studies highlight the potential of unsupervised learning techniques to improve the accuracy and effectiveness of breast cancer diagnosis by histological imaging of this deadly disease.

2.3.5. Diagnosis of histological images of breast cancer patients using integrating multimedia data techniques

There have been some recent studies that have led to significant improvement in breast cancer diagnosis by analyzing breast cancer tissue images as a result of integrating multimodal data fusion techniques. The most prominent of these studies are the following:

1. Abdullakutty et al. [44] conducted an extensive review on improving histopathology-based breast cancer diagnosis, highlighting the integration of multi-modality data and the significance of interpretability in AI models.
2. To enhance diagnostic accuracy across diverse staining methods, Modi et al. [45] suggested a multi-stained, multi-level zigzag network to differentiate various tissue types in breast cancer images.
3. To improve the classification of high-resolution histology images, Zhong [46] constructed a deep spatial fusion network, which in turn demonstrated improved accuracy in distinguishing carcinoma from non-carcinoma tissues.
4. for the segmentation of multi-class breast cancer images, Ho et al. [47] presented deep multi-magnification networks, employing data from different magnifications to enhance segmentation precision.

The entire above are summarized and the details and results of these studies are combined for breast cancer classification methods and presented in Table 1.

This table categorizes the approaches according to their contributions to pattern recognition, diagnostic

Table 1. Summarizes the techniques used to understand and distinguish patterns of breast cancer, achieve greater diagnostic accuracy, and improve treatment methods.

Technique	Goal/Purpose	Models	Year
Machine Learning (ML) Algorithms	Its purpose is to recognize patterns and improve classification.	K-NN, RF, NB, and SVM; focus on shape, texture, and density [29].	1990–2001
Deep Learning (DL)	Automate the process of feature extraction from datasets and improve scalability.	CNNs like AlexNet and VGGNet applied to breast cancer histology [31, 48].	2016–2018
Hybrid Architectures (CNN + RNN)	Their purpose is to integrate spatial and sequential features that enhance the accuracy and effectiveness of advanced diagnostic techniques	CNN-RNN combinations for capturing diverse patterns [34].	2018
Generative Adversarial Networks (GANs)	Enhance datasets to tackle issues associated with limited data and imbalance.	Synthetic data generation for classification improvement [35].	2018–2020
Transfer Learning (Dual Transfer)	Purpose to achieve better classification, pre-trained models were utilized.	Training on unclassified data followed by labeled datasets [23].	2023
Lightweight CNN Models	Its purpose is to make optimal use of resources within limited environments.	Reducing computational costs without sacrificing accuracy [36].	2023
Multimodal Data Fusion	Its purpose was to integrate diverse medical data sources to obtain a comprehensive diagnosis.	Combining imaging with chromatin and genetic data. [21, 44]	2020–2024
Unsupervised Learning Techniques	Their purpose is to identify patterns in data without labeled data.	Clustering, autoencoders, and GANs applied to mammograms and histopathology images [41, 42]	2018–2023
Invariant Supervised Self-Learning	Its purpose is self-supervised learning is to enhance resilience to color variations.	Stain-invariant self-supervised techniques for histopathology analysis [42].	2022
Deep Multi-Magnification Networks	Its purpose is to use multiple amplifications (zooms) to improve segmentation.	Enhancing segmentation through diverse magnification inputs [47].	2019

accuracy, and treatment improvement. It is necessary to understand the nuances within the different histological images in order to devise focused treatment methods. By finding connections between visualization features and therapeutic responses, we can enhance treatments and improve patient care [49].

3. Overview of the 2018 ICIAR database of histological images

The 2018 ICIAR2018 BACH Conference on Image Analysis and Recognition was created in 2018 to challenge breast cancer tissue image data, with the goal of providing computer scientists and clinicians with tools for breast cancer diagnosis [50]. This collection includes a large variety of images, showing different types and stages of cancer tumors. Techniques and algorithms for classifying breast cancer are created and evaluated using this database [51].

ICIAR is known for its support for researchers and scientists in image classification because it promotes research and innovation in medical image analysis, and is a great force and a useful tool for them. This database was launched in 2018 to classify histological images of breast tumors. A database of cancer tissue images was launched before ICIAR, and it is a useful tool for researchers in the field of classifying histological images of tumors [43].

Malignant tissue images classified as objective, invasive, and benign. There are 2042 images in the ICIAR2018 database. Through the various categories in this database, it can represent breast cancer of different types and stages. This collection contains images from multiple sources representing a wide variety of tissue types, which shows the quality of the images in it [51].

Using the ICIAR 2018 database, researchers can develop and test machine learning and deep learning models, CAD tools, and algorithms for breast cancer classification. This vast amount of information has allowed researchers to train their algorithms on a variety of cases, helping to develop their classification techniques. By making it easier to compare different algorithms, this dataset enables advances and developments in the industry [52].

In general, patients have become more aware about their treatment plan options through the use of histological images in staging breast cancer. The availability of data sets such as the R&D Efforts Database has contributed to the advancement of this field. To achieve the real goal of more accurate patient outcomes, researchers are working to increase the efficiency and accuracy of breast cancer screening using advanced technology and large amounts of data [53].

4. Breast cancer classification and addressing its obstacles and challenges

Despite the progress achieved by histological imaging in staging breast cancer, there are several challenges and limitations that need to be addressed. The diversity of complex tumors, discrepancies between medical observers, different image quality, and lack of standardization in staging criteria for complex tumors are some of these challenges and limitations [54]. If we wish to improve the quality of patient care and outcomes by adopting an advanced approach to accurate and standardized classification [55].

We must overcome these limitations and challenges if we wish to improve the quality and outcomes of patient care by adopting an advanced approach to accurate and standardized classification [56].

We look forward to the future that future advanced technology will contribute to improving breast cancer detection techniques due to the continuous and rapid technological progress and innovation in addition to the collaborative efforts to care for breast cancer patients [57].

4.1. A selection and extraction of characteristics

The process of identifying and extracting important features is important for classifying histological images of breast tumors. These features provide important details that can accurately identify the type and stage of disease in the current case. The ICIAR 2018 database was relied upon with the help of computer to extract these features by applying different methods and selecting the best one that has high results to be used in classifying the data [53].

4.2. Classification of breast cancer histological images by using common features

Researchers found common features in histological images that are commonly used to classify breast cancer. Among these features are:

Unique morphological features: These features are used to record the special characteristics of cancer cells, including the size, shape, and texture of tissues and cells that are visible in histological images. They provide important information about the composition of cancer cells, which helps researchers and medical experts distinguish between types and stages of breast cancer by providing important details regarding the structure of cancer cells [58].

Features of statistics: One of the statistical features in histological images is the relationship between pixel intensity and distribution analysis. These features can aid in classification by revealing the texture, density, and spatial arrangement of cells [59].

Features of Texture: Textural features of tissue are provided by histological images. They have the ability to record variations and patterns that may indicate certain types of cancer. Wavelet transform-based features, gray range length matrices, and co-occurrence matrices are all examples of such common synthetic features [60].

4.3. Techniques for selecting ideal features for breast cancer classification

For accurate breast cancer classification, it is just as vital to extract the right features as it is to extract all of the features. Feature selection is done using the following methods:

4.3.1. Principal component analysis (PCA)

One dimensionality reduction technique is principle component analysis (PCA), which seeks to convert the original features into a new collection of independent variables known as principle components. These primary components allow for the identification of the most useful features by capturing the largest variance in the data as show in Fig. 5 [61, 62].

4.3.2. Recursive feature elimination (RFE)

Beginning with all of the features, RFE iteratively removes the least significant ones according to a predetermined criterion as show in Fig. 6. This process is carried out until either the required number of features or some other predetermined endpoint is met [63, 64].

4.3.3. Genetic algorithms (GA)

It is a method of optimization and research. This method can be classified as one of the methods of evolutionary algorithms that rely on imitating the work of nature from a Darwinian perspective. A genetic algorithm uses a search technique to find exact or approximate optimal solutions. It involves creating a set of subsets of potential features, evaluating each one against an objective function to ensure its fit, and then gradually improving the set using genetic

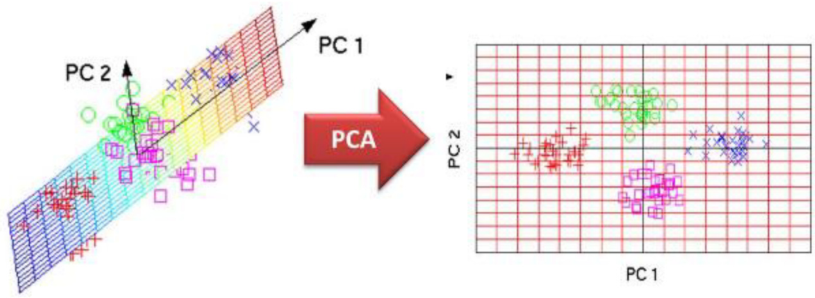

Fig. 5. Illustration of Principal Component Analysis (PCA) in ML [62].

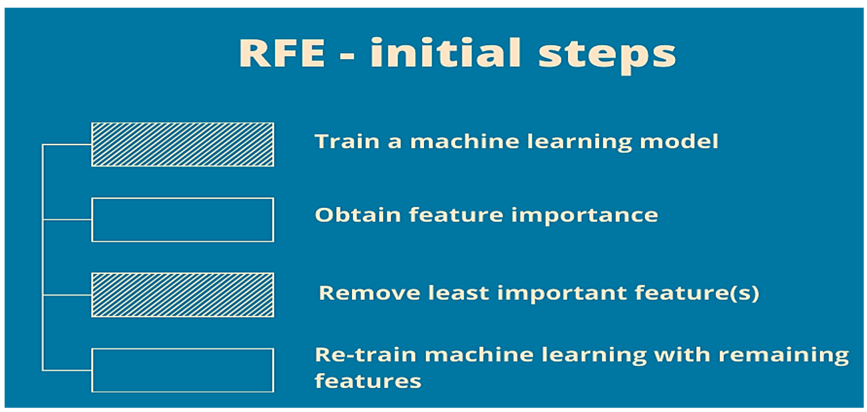

Fig. 6. Explain initial steps of RFE [64].

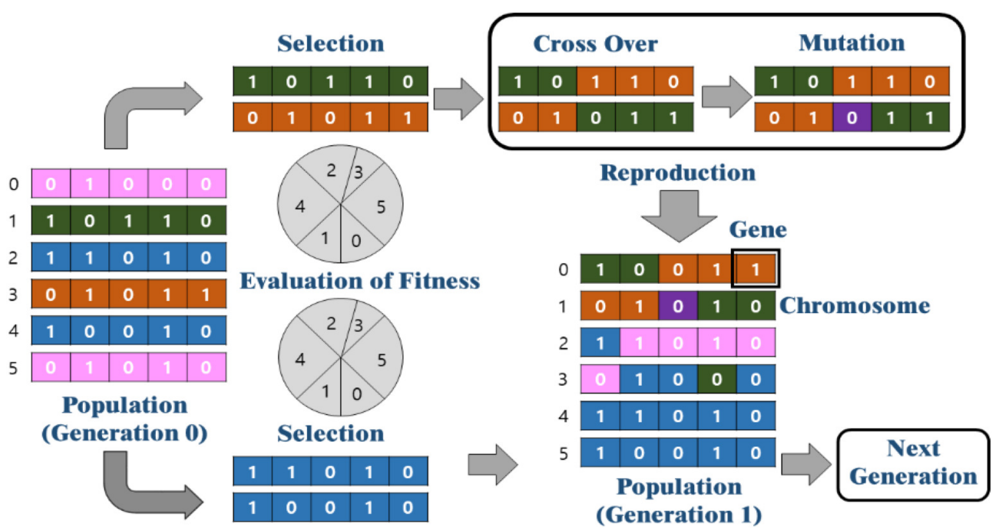

Fig. 7. Genetic selection method using genetic algorithm [66].

operators such as crossover and mutation as show in Fig. 7 [65].

4.3.4. Encapsulation (Wrapper) methods

The encapsulation method adds feature selection to the classifier training process. They test multiple subsets of features for predictive performance by the chosen classifier and choose the classifier that performs best. In integrative methods, external information is used as a filter to reduce the feature search space before passing the reduced feature set through a wrapper or inline method to obtain the final feature set as show in Fig. 8 [67].

Studies have shown that these feature extraction and selection methods enable them to efficiently classify breast cancer histological images. These techniques further improve classification accuracy, as

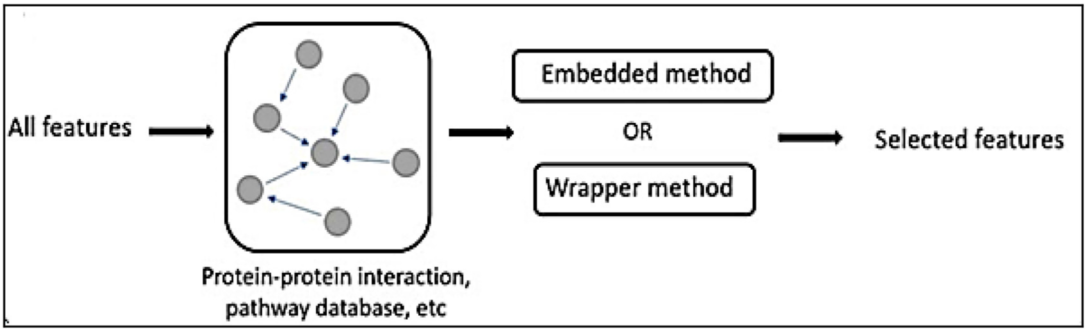

Fig. 8. Encapsulation (Wrapper) methods [67].

Table 2. Approach and methodology for classification.

Implementing Preprocessing Methods	Feature Extraction	Classification Algorithms
Some pre-processing techniques (noise reduction, image normalization, and color correction) are important and necessary techniques that are used before classification, which work to improve image quality and reduce the effect of different acquisition conditions [68].	Various algorithms are used to extract relevant features from histological images, including Gabor filters, texture analysis methods, and wavelet transform. These features help distinguish between healthy and malignant tissue by picking up distinct patterns [69].	Breast cancer histological images are categorized using a variety of classification algorithms, such as SVMs, neural networks, and RFs. Using the collected features, these algorithms classify images accordingly [70].

Table 3. The results and the analysis.

Overall Accuracy	Best Class of Performers	Class with the Poorest Results
SVM, which stands for support vector machines, was the approach that performed the best overall. The greatest accuracy that was reached was 86.49%.	With an accuracy rate of 91.17 percent, the invasive carcinoma class was the simplest to classify.	With an accuracy percentage of just 70.06%, normal breast tissue proved to be the most challenging to classify.

Table 4. Limitations of ICIAR database 2018.

Small Dataset	Noisy Labels	Limited Information
The ICIAR database 2018 contains only 2,000 images, which may not be sufficient to train DL models effectively.	It can be difficult to use a dataset for research because the labels that come with it are not always reliable.	The utility of the dataset may be limited because it only includes tissue images and does not include any clinical information about patients.

well as our understanding of the disease, and help find more targeted treatments.

Obtaining an optimal classification is very important, and this is achieved by extracting and selecting features from breast cancer histological images. These features, coupled with modern machine learning algorithms, can enable healthcare to identify and treat breast cancer very effectively, as shown in Tables 2 to 4 [51].

4.4. Breast cancer histological image classification challenges

Pathologists and medical researchers face several challenges in identifying histological images of malignant diseases such as breast cancer. A good understanding of these challenges and difficulties is essential to developing effective and reliable classification models. Let's look at the key points surrounding this subject:

4.4.1. Variation in tissue preparation and coloring appearance

Variability in tissue preparation is a major challenge in classifying histological images of breast cancer. The appearance and quality of images may vary based on the methods used by technicians in different laboratories. This may result in inconsistent diagnoses and the development of standardized classification criteria [71].

4.4.2. Complexity and diversity of textile patterns

Breast cancer is often observed to have unimaginably complex and heterogeneous histological manifestations. Diverse tissue compositions characterize different breast cancer subtypes, such as invasive ductal carcinoma (IDC) and ductal carcinoma in situ

Table 5. The breast cancer histological image classification challenges.

Challenge	Description
Variability among observers	Determining the features of breast cancer histological images is a task fraught with considerable intra- and inter-observer heterogeneity.
Noise in the background	What makes it very difficult to classify cancer cells is that medical images contain a large amount of noise.
An imbalance in the data	With a far higher proportion of benign images than malignant images, the ICIAR database features an imbalanced class distribution.

(DCIS). Within each classification, slight variations can also exist in the structure and arrangement of cells and tissues. Because of these many differences, it can be difficult to differentiate between different types and subtypes of breast cancer [27].

4.4.3. Overlapping and ambiguous features

The presence of complex and ambiguous features is another challenge in tissue classification in breast cancer images. Some histologic features may be difficult to define precisely because of their similarity to other breast cancers. A need for unbiased and reliable classification methods arises when pathologists disagree in diagnosis [72].

The Breast Cancer Histological Image Classification Challenges shown in Table 5.

Classification of breast cancer histological images faces several challenges. Accurately understanding these images is difficult due to their complexity and diversity in histological patterns, differences in tissue preparation and staining, in addition to the presence of overlapping and ambiguous features. To improve the accuracy of breast cancer diagnosis, researchers and histopathologists are working hard to develop accurate and reliable classification algorithms.

5. Breast cancer histological image classification limitations

To correctly detect breast cancer, correct interpretation of histological images is essential in this complex process. Although the results of classifying these images using artificial intelligence and machine learning algorithms have been promising, there are limitations and difficulties to consider [73].

5.1. Limited access to training data with annotations

One of the main difficulties in classifying breast cancer histopathological images is the lack of reasoned data for training. Reported data is key to teaching AI systems how to accurately recognize and classify different forms of breast cancer. But obtaining a large set of well-tagged images is a tedious and time-consuming process. The limited amount of annotated data may have an impact on the models' ability to classify accurately and how well they generalize [74].

5.2. Challenging in capturing spatial information

Another difficulty is the issue of obtaining spatial data from tissue images. The term “heterogeneity” conveys the idea that different features and patterns are present in different regions of breast cancer tissue samples. Obtaining and analyzing spatial information appropriately is vital for correct classification. But even with the current limitations of how classification methods do not effectively capture or use this specific information (despite their best efforts), misclassification can still occur [53].

5.3. Resolution and magnification effects on images

Creating histological images at different resolutions and magnifications can also present additional challenges for classification because they can be difficult to resolve. Different imaging modalities and techniques are likely to produce images at different levels of resolution and magnification. If a particular feature is useful in making a correct classification but its salience is compromised due to contrast and due to image quality, it can be very difficult for the model to make accurate predictions [75].

Although leveraging artificial intelligence and machine learning algorithms to classify breast cancer histological images is already showing good results, it is far from an easy process. Some typical barriers include: lack of an easy way to obtain spatial data; Lack of sufficient explanatory data; and many other factors such as image resolution and zoom can affect the final results. However, despite all these challenges, we have yet to be able to formulate a robust classification model that will play an important role in detecting breast cancer in its early stages with high accuracy.

6. Challenges and limitations: solutions (methods and techniques)

6.1. Uniformity in the methods used for staining and preparing tissues

One challenge in identifying images of breast cancer tissue is that tissue preparation and staining procedures are not standardized. Variations in laboratory techniques can affect the appearance of tissue samples, making it difficult to compare and classify images successfully. Therefore, researchers are actively working to standardize staining and tissue preparation processes. To enable accurate comparisons across laboratories and classification of breast cancer-related histological images [76].

6.2. Development of algorithms for image processing and analysis

Classification is challenging due to the complex structure of breast cancer histological images. Advanced algorithms can improve breast cancer classification by detecting and analyzing complex and difficult patterns and structures that may be difficult to resolve visually [77].

6.3. Multimedia data integration for taxonomy development

Researchers aim to improve breast cancer image classification based on tissue composition by combining genetic, clinical, and histological data. Combining diverse data sources is essential for a comprehensive understanding of cancer cell characteristics. Combining data from multiple sources creates more accurate breast cancer models that can differentiate between cancer types and provide personalized treatment recommendations [78].

Improving multi-tissue classification of breast cancer images involves standardization, advanced image processing, and combining multiple data sources in a single location. These techniques aim to increase accuracy and efficiency and overcome the obstacles and limitations associated with classification to improve patient outcomes.

6.4. Difficulties and challenges in breast cancer imaging utilizing deep learning

There are several challenges and obstacles that hinder the effective implementation of DL in breast cancer imaging accurately despite the advances in technology, including:

1. Limited data and labeling: Labeled datasets are very important for training deep learning models and helping them understand and classify the information they process. Acquiring labeled pathological tissue images is highly labor-intensive and requires specialized knowledge from pathologists [51].
2. Variation in data quality: Differences in imaging equipment, differences in preparation methods, and variation in tissue staining procedures can introduce noise that ultimately affects the efficiency of model performance [43].
3. Class imbalance: Most datasets have an irregular distribution of data between classes, for example, the proportion of benign cases is higher than that of malignant cases, which leads to biased predictions of benign cases and thus lowers accuracy in detecting malignant cases [37].
4. Large computational requirements: Deep learning models require large computational resources, which limits their deployment in resource-limited environments [54].
5. Lack of standardization: Variation in preprocessing methods and assessment measures complicates comparability and reproducibility between studies [58].

Solutions to the above challenges through the use of deep learning techniques in breast cancer imaging are as described below:

1. **Limited data and labeling**:
   - Data augmentation is a method of increasing the size and diversity of deep learning training datasets. It involves applying random changes to existing data, such as flipping, cropping, rotating, and changing colors, which may artificially increase the size of the data set [79].
   - Synthetic Data Generation: Using Generative Adversarial Networks (GANs) to create real-world synthetic texture images [38].
   - Transfer learning allows developers to leverage knowledge gained from pre-trained models to solving one problem and apply it to a different, but related problem [80].
   - Active Learning: The main idea is to select only the most informative samples with the help of pathologists to annotate them and then train a model with these samples in a supervised manner [81].
2. **Variation in Data Quality**:
   - Stain Normalization: Stain normalization is an innovative solution to overcome stain variability by Standardizing tissue staining across images, and has been commonly used in

computer-aided diagnosis (CAD) systems using computational algorithms [82].
  - ○ Quality Control Pipelines: Medical image preprocessing is a critical step to ensure uniformity in image resolution and staining by improving image quality [83].
  - ○ Domain Adaptation: Domain adaptation deals with the problem of adapting a model trained on data from one domain (the source domain) to perform well on data from another domain (the target domain) [84].
3. **Class Imbalance**:
   - ○ Resampling techniques: To rebalance the class distribution of an imbalanced data set, resampling techniques provide over-sampling of malignant cases or under-sampling of benign cases [85].
   - ○ Cost-sensitive learning: Assigning higher penalties to instances misclassified as malicious, thus reducing the cost of misclassification of the model on the input data [86].
   - ○ Synthetic Oversampling (SMOTE): The component works by creating new synthetic cases for underrepresented classes from existing minority cases that you provide as input [87].
4. **Large computational requirements**:
   - ○ Model Optimization: Using lightweight architectures like MobileNet for resource-constrained environments, achieving a streamlined transformation of deep learning models [88].
   - ○ Edge computing is a distributed computing framework that brings enterprise applications closer to data sources such as IoT devices or local edge servers [89].
   - ○ Federated learning: Distributed training on heterogeneous datasets across multiple machines to ease the computational load [90].
5. **Lack of Standardization**:
   - ○ Benchmark datasets: An algorithm based on a standardized dataset to effectively reduce variance, and jointly develop standardized datasets with uniform preprocessing procedures [91].
   - ○ Evaluation frameworks: Adopting standardized metrics such as F1 score is based on a specific threshold, while area under the curve (AUC) evaluates all thresholds, providing insights into different classification aspects of model evaluation [92].
   - ○ Community Efforts: To enhance reproducibility, it is best to share open source code and models with a broader community that develops them together [93].

The following detailed solutions are proposed to address the obstacles and challenges in preprocessing in histopathological image analysis, including variations in staining and sample preparation:

1. **Stain Normalization Techniques**:
   - ○ To preserve the structure and uniformity of color differences across texture images it is best to use “stain normalization”. This improves consistency between datasets [82].
   - ○ One of the stain normalization techniques that enhances data normalization is the use of “GAN-based methods” that produce synthetic images [38].
2. **Quality Control Pipelines**:
   - ○ To reduce discrepancies in imaging results, it is preferable to implement automated pipelines to check the consistency of staining, image resolution, and noise levels before analysis [83].
3. **Advanced Imaging Equipment**:
   - ○ To ensure that the results of imaging samples are uniform, it is preferable to use scanning devices that are equipped with built-in correction algorithms for spots and inconsistencies in imaging and are also of high resolution [54].
4. **Standardization of Sample Preparation Protocols**:
   - ○ To ensure reproducibility and consistency of diagnostic results, it is best to adopt universal guidelines for tissue fixation and staining to minimize interlaboratory variability. These protocols for sample preparation include: consistent tissue fixation, standardized sectioning, and controlled imaging conditions [71].
5. **Data Augmentation**:
   - ○ Data augmentation is a technique to increase the number of labeled examples required for deep learning training. It artificially enlarges the original training dataset by introducing various transformations such as translation, rotation, scaling, and even noise, in addition to using test time augmentation (TTA), which involves applying transformations such as horizontal and vertical flipping to the original data instances to create new instances [94].
6. **Integration of Genomic and Clinical Data**:
   - ○ For consistent analysis, deep learning algorithms have automated the segmentation of epithelial and histological tissue in images while integrating imaging and genomic and clinical data to reveal relationships, helping to understand breast cancer and improve treatment strategies [21].

7. Opportunities for development and future directions

7.1. Enhancing classification accuracy with DL methods

The use of artificial intelligence in classifying breast cancer tissue images from the 2018 ICIAR database shows promising potential due to the advancement in the field of machine learning and image analysis. Some problems related to image classification have shown encouraging results for DL models, namely convolutional neural networks (CNNs). Using these algorithms to improve the quality of histological image classification could help researchers diagnose breast cancer more accurately [53].

DL, a branch of ML, has contributed greatly to medical image classification, especially classification of breast cancer histological images [95].

CNN is the most advanced algorithm currently available. These trends can automatically learn and record distinct patterns from images, meaning there is no need to put effort into manually extracting features. A CNN is based on many interconnected neurons and is able to discriminate within important elements and patterns of images [96]. ResNet, AlexNet, GoogLeNet, and VGGNet are the most commonly used CNN architectures for breast cancer classification. When used on histological images, these models showed complete accuracy in distinguishing between cancerous and benign cells. These models have been fine-tuned for breast cancer classification using transfer learning, a method that allows pre-trained models to be fine-tuned to specific tasks [97]. Creating synthetic tissue images using generative adversarial networks (GANs) has shown promise as a way to augment limited datasets and address data imbalance issues recently [98].

Combining the use of ML algorithms with DL models has significantly improved the classification of breast cancer histological images. These newer methods are expected to support pathologists in more accurate diagnoses, which could be beneficial for breast cancer patients [12].

7.2. Enhancing classification accuracy with transfer learning methods

Transfer learning addresses the challenges posed by small labeled datasets, emerging as a transformative approach to breast cancer classification. Researchers can fine-tune these infrastructures for specific tasks. By leveraging pre-trained models, such as ResNet, Xception, and VGGNet, which are initially trained on large visual database sets designed for use in optical object recognition software research such as ImageNet, this in turn reduces overfitting as well as the need for a dataset. Classified extensively and thus obtain robust models as quickly as possible [23].

In another study, Makhliv et al. [23] proposed a dual transfer learning strategy. This method involves training in two parts, first on unlabeled medical images and second fine-tuning on labeled datasets. The advantages of this approach were improved classification accuracy and enhanced the ability of models to adapt to variations such as coloration and resolution in image datasets.

7.3. Enhancing classification accuracy with unsupervised learning methods

A major opportunity for unsupervised learning techniques is the "abundance of unlabeled histopathological images," which can extract patterns of interest without having to refer to labeled data. Clustering, self-supervised learning, and feature discovery are key features that can be leveraged from this untapped resource [22].

Xiao and Lu [22] improved classification by combining unsupervised learning with semi-supervised methods. Their method significantly improved classification performance by clustering similar features and propagating labels from a small labeled subset to unlabeled images. This approach proved effective for groups Large-scale breast cancer data by addressing misalignment between categories and variation in image quality.

7.4. Enhancing classification accuracy with multimodal data fusion methods

Multimodal data fusion represents a promising prospect in breast cancer diagnosis, as it integrates various sources of information such as histopathological images, clinical records, and genomic data. Researchers can achieve a more comprehensive understanding of tumor biology and improve the accuracy of diagnosis by combining these methods [21].

S. Xu et al. [21] in their study highlight the impact of combining genetic information and histological imaging. This integration allowed for deeper analysis of estrogen receptor-positive breast cancers, revealing patterns that were not previously detectable through single-pattern approaches. Such advances pave the way for personalized diagnostics and targeted therapies.

7.5. Developing platform for collaboration in data sharing and benchmarking

To break down the barriers imposed on the 2018 ICIAR database and encourage further research into breast cancer staging, this is done by adopting collaborative systems to facilitate data exchange and measurement. These systems can increase the diversity and quantity of datasets available for training and evaluation by supporting researchers at different universities with their own data input and algorithms. When researchers collaborate to achieve a common goal, they can expand and enhance the generalizability of taxonomic models, accelerate the development process as well as enhance teamwork [99].

7.6. Integration of genetic and clinical data with the purpose of conducting extensive analysis

While the 2018 ICIAR database contains histological images useful for research, it should be noted that visualization alone may not provide a complete understanding of breast cancer. Integrating genomic information with clinical information helps provide a more complete analysis of breast cancer subtypes and corresponding images, such as gene expression patterns and mutational data, along with demographics, tumor stage, and treatment history of the patient. Consolidating additional data sources can help researchers discover more precise analyses, discover new biomarkers, and improve personalized treatment regimens [21].

The challenges and limitations of classification of breast cancer histology images from the 2018 ICIAR database give an opportunity for further research and progress. By demonstrating DL models, creating collaborative platforms for sharing and measuring data, and integrating genomic and clinical data, researchers can enhance the quality of breast cancer detection and treatment.

7.7. Proposed solutions and future directions

- Data augmentation: While operations such as translation, rotation, flipping, scaling, and cropping increase the diversity of datasets, GANs solve class mismatch issues by creating synthetic models [94].
- Learning transfer: Leverage pre-trained models like ResNet and Xception to reduce reliance on large data sets and speed up model adaptation for specific tasks [78].
- Efficient model performance: CNN is lightweight, and adaptive anti-aliasing reduces computational complexity, providing real-time analysis [97].
- Multimodal data fusion: Combining histopathological images with clinical and genetic data improves the accuracy of diagnosis and provides comprehensive information on different types of cancer [21].

Therefore, microscopy can play a pivotal role in early detection and treatment of breast cancer, improving patient outcomes and reducing diagnostic errors by addressing these challenges and adopting advanced technologies.

In Table 6 below, we compare the strengths and weaknesses of the techniques used in the ICIAR 2018 breast cancer classification tasks as well as the problems addressed along with the objectives and importance.

Table 6 highlights a diverse range of methods utilized in the ICIAR 2018 breast cancer classification task, each with distinct advantages and limitations. The strengths of these techniques include high classification accuracy e.g., Minh et al. [100], innovative hybrid architectures (e.g., Yan et al. [34]), and computational efficiency e.g., KAUSA et al. [36]. However, many methods face challenges such as reliance on extensive labeled datasets, loss of critical image features during preprocessing (e.g., Nassima et al. [35]), and sensitivity to model complexity (e.g., Guo et al. [101]). These findings underline the need for approaches that balance accuracy, interpretability, and computational feasibility to address the multifaceted challenges of breast cancer classification.

From the results presented in Table 6, an effective approach to balancing accuracy, interpretability, and computational feasibility for breast cancer classification involves integrating several strategies:

Hybrid Modeling Techniques: Combining different architectures, such as convolutional and recurrent neural networks (e.g., Yan et al. [37]), captures both local and global image features while preserving spatial relationships. This approach enhances accuracy and provides interpretable insights into feature importance.

Transfer learning and data augmentation: Leveraging data augmentation and transfer learning (e.g., Nguyen et al. [94] and Mukhlef et al. [23]) helps address the challenges of limited datasets by improving model generalization without requiring large computational requirements. It also allows models to adapt to varying shooting environments.

Low-cost architecture: Using efficient network designs, such as the Lightweight Deep CNN model with Wavelet Transform (e.g., KAUSA et al. [36]), reduces memory and computation costs while maintaining

Table 6. Techniques used in the breast cancer classification tasks for ICIAR 2018.

Authors	Methods/ Techniques	Strengths	Weaknesses	Problems Addressed	Objectives	Importance
Minh et al. [100]	In this study, convolutional neural networks (CNNs) were used to extract the features and then classify them. To adjust the pre-trained CNN models to medical imaging tasks, transfer learning was used. To deal with the limited dataset sizes, data augmentation was implemented.	As for the power of this technique, the classification accuracy was high at 87.5% for distinguishing between cancer and non-cancerous and 80% for four cancer classes. In addition to the effective use of data augmentation for training power, the reliability of medical diagnosis was improved by automating image analysis.	The research is in its infancy although there is room for improvement in CNN structures and methodologies. Apart from relying only on labeled datasets, several important factors that affect classification performance should also be considered.	Problem solved: The problem of manual analysis comes from the size and complexity of the data, reducing the risk of erroneous analysis and time-consuming processes.	Develop automated methods to classify breast cancer images and differentiate between cancer types to improve clinical decision-making.	The importance of the technology lies in reducing human error and the workload on the pathologist, and improving early diagnosis and treatment of patients.
Guo et al. [101]	They propose a high-performance, scalable architecture to address. The four-category problem of breast cancer based on deep CNN networks.	Our hybrid CNN unit might use image local and global information to produce more accurate predictions. Bagging and hierarchical voting are also used to improve the classifier. In terms of total accuracy and sensitivity in classifying each category of images, our approach outperforms the state of the art, according to trials.	The structure of the model, particularly the GoogleNet model, is quite complicated.	This approach addresses the challenges inherent in breast cancer history, minimizing errors, managing small data sets, and integrating national and international data to ensure accurate predictions from the data.	This technology addresses the challenges of automating breast cancer tissue image recognition, reducing error rates, and dealing with limited data. The technology achieves highly accurate predictions by integrating global and local information.	The importance lies in automating the diagnostic process, which in turn reduces the workload on pathologists, improves accuracy, and improves early detection, which is very important to reduce mortality in breast cancer cases.
Anupama et al. [102]	Using histology images, this paper classifies breast cancer kinds. Histology images can be classified using image processing. The current study uses capsule networks to record spatial and orientation data.	This paper shows that data pre-processing and parameter adjustment increase traditional architecture performance. The results demonstrate that this technology can be used as an automated tool to help clinicians diagnose diseases, which may boost cancer survival by focusing on treatment rather than diagnosis.	The process of extracting 10 patches from each image may cause the loss of some important features that are relied upon to determine the type of disease.	The problems I addressed were that traditional CNN architectures for classifying breast cancer tissue images failed to capture directional and spatial relationships in the data. There is also a need to improve classification accuracy with the limited amount of data that exists.	Its goal is to improve both accuracy and computational efficiency and is done by enhancing the classification of breast cancer tissue images using capsule networks with patch extraction pre-processing and macular normalization techniques.	This technology provides powerful automated tools for doctors to help them improve the quality of diagnosis, make early and accurate diagnoses, and focus on treatment rather than diagnosis, which greatly improves survival rates.

(Continued)

Table 6. Continued.

Authors	Methods/ Techniques	Strengths	Weaknesses	Problems Addressed	Objectives	Importance
Kassani et al. [103]	A DCNN descriptor and pooling operation enable DL-based breast cancer categorization. We employed data augmentation methods to improve categorization. Various stain normalization methods are also tested.	Experimental results show that the pre-trained Xception model-based DCNN architecture exceeds all others in classification accuracy with 92.50%.	The work needs to use DL based ensemble models and improved speckle normalization to increase the classification accuracy.	The problem of color contrast in color images and limited data size were addressed, both of which are important for early and accurate detection of breast cancer using histopathological images.	Develop a deep learning-based framework that uses transformative learning, global integration, and data augmentation to classify breast cancer images into 4 groups: “normal”, “benign”, “In situ”, and “invasive”.	This approach helps in early and accurate diagnosis of cancer, reduces the workload of pathologists, and improves patient outcomes in their survival rate through timely and focused care.
Nguyen et al. [94]	This technique improves histology image classification by increasing the dataset (400 to 480) images through additive patch extraction(APE) and data augmentation, using Test time augmentation (TTA) during testing with CNNs.	78% accuracy on the four-class test set prediction was attained, which is a respectable result when compared to the 65% accuracy rates that were previously published in Vu et al publication. The evaluation includes five-fold cross-validation, which improves the classification accuracy into four cancer classes.	It would be beneficial to further optimize the model during training and expand the study to include other data sets or domains.	The problem of the limited accuracy of deep learning models is solved by solving the problem of the scarcity of labeled data for breast cancer tissue image classification.	The objective was to improve the dataset by improving classification accuracy using a time-based test (TTA) approach and removing some spots from the whole-slide image (WSI).	Improves diagnostic efficiency by automating histology image classification, reduces subjectivity, and addresses data limitations critical to accurate cancer detection.
Yan et al. [37]	A hybrid CNNs and RNNs for breast cancer histopathology image categorization is proposed. Our method combines convolutional and recurrent neural networks with the deeper multilevel feature representation of histopatho-logical image patches to retain short-term and long-term spatial relationships.	With an average accuracy of 91.3% for the 4-class classification test, our system exceeds the state-of-the-art method, according to the experimental data.	It is ideal to directly use a complete high-resolution image as input to a deep neural network.	Issues identified are the importance and complexity of analyzing histopathological images, lack of definitive data, and challenges associated with processing high-resolution histopathological images of breast cancer.	The objective is to improve automated breast cancer image classification. A hybrid model using CNNs and RNNs captures spatial features and context, enhancing classification performance for better diagnostics.	The importance of the technology is highlighted by automating image classification and providing a reference data set to guide progress in medical imaging and reducing reliance on medical experts, all of which have helped in the accuracy of medical diagnosis.

(Continued)

Table 6. Continued.

Authors	Methods/ Techniques	Strengths	Weaknesses	Problems Addressed	Objectives	Importance
Nassima et al. [35]	This article makes use of the Inception -v3CNN architecture, six histopathology source datasets, and four target sets as its foundational components. To address class imbalance in histopathology datasets, transfer learning and GANs have been used.	The proposed strategy outperformed the alternatives in the literature on the CRC (95.28%) and KIMIA-PATH (98.18%) datasets, according to the comparison research.	The process of cutting the image into patches may lose some of its properties, i.e. there may be important features that may be lost. Instead, they can use image resizing and data augmentation to increase the number of samples.	The problem addressed was the challenges in transferring knowledge from large public datasets (e.g., ImageNet) to specialized histopathological tasks, as well as the scarcity of annotated data for medical image analysis.	Its objective was to use models pre-trained on similar tasks to improve classification accuracy and demonstrate the effectiveness of intra-domain transfer learning between histopathological datasets.	The importance of this technique is highlighted in improving the efficiency of transfer learning in medical imaging, which has improved performance in pathological image classification tasks and reduced dependence on distant datasets.
Rafael et al. [104]	Using three classifiers for automatic breast lesion classification in medical images and Transfer Learning for feature extraction, this work compared and evaluated the performance of four network designs.	This work's best algorithm was ResNet50, which used the SVM classifier with Polynomial kernel to get values above 78% in evaluation metrics. Good evaluation findings allow breast lesion research goals to be proposed.	They did not use other CNN architectures like Transfer Learning to compare the results or other classifiers.	The problem of difficulty in accurately classifying breast lesions due to limited data,potential errors in manual analysis by specialists, and high computational cost was solved.	To evolve an methodical, low-cost computational procedure for automatic classification of breast lesions using transfer learning with classifiers and multiple CNN architectures.	It is important in providing a second expert opinion for diagnosis, helps in early detection of breast cancer, serves as a reliable tool for CAD procedures in clinical practice, and reduces false posi-tive/negative results.
Amr et al. [105]	An automated classification method that combines multilayer hand-crafted features with pre-trained deep CNNs to extract features.	When DL features were combined with handcrafted features, the performance classification results improved to 96.79 percent for four-class classification.	They did not check the classification performance when training the first layers of the pre-trained models on tissue images similar to breast cancer tissue images such as colon and bone cancer tissue images instead of keeping the first layers trained on ImageNet.	It addressed the problem of the limitation of hand-crafted features or deep learning in breast cancer histopathological image classification and the challenges in dealing with the limited dataset size for medical image analysis, which may affect its generalizability and clinical application.	Its objective is to combine both the features of deep learning and manual learning for multi-class classification of histopathology images, and also by using transfer learning techniques, this leads to overcoming the limitations of the dataset and thus improving the classification performance.	It significantly improves classification accuracy by leveraging the complementary strengths of individual hand-crafted or deep learning features in classifying breast cancer histopathology images, which in turn has helped in better diagnosis and provided a powerful CAD tool for breast cancer detection.

(Continued)

Table 6. Continued.

Authors	Methods/ Techniques	Strengths	Weaknesses	Problems Addressed	Objectives	Importance
KAUSA et al. [36]	Wavelet Transform (WT) and Lightweight Deep CNN Model	Reduces computational cost by focusing only on low-frequency bands of images. Designed with invertible residual block modules, which reduce memory usage and computational complexity.	Some image information may be lost due to decomposition, which might affect detailed texture analysis. May not be as accurate for complex datasets without careful opti	We address the problem of high computational cost and memory requirements of deep CNNs for classifying breast cancer pathological images, which makes them time-consuming and unsuitable for capacity-constrained platforms.	Its objective is to achieve efficient multi-class classification of breast cancer histological images by designing a lightweight CNN model that leverages WT and invertible residual blocks.	Its importance lies in the fact that it significantly reduces computational costs without compromising classification accuracy, and thus can be used immediately and helps medical professionals detect cancer early.
Xiao et al. [22]	Semi-Supervised Classification	Combines labeled and unlabeled data to improve model performance; spreads label information to unlabeled data.	Clustering can be sensitive to the initial data distribution and may not always find meaningful patterns in complex data.	It addressed the problem of the difficulty of classifying large medical images due to its reliance on expert knowledge as well as taking advantage of large-scale, unlabeled data to classify medical images effectively.	Their objective was to develop a semi-supervised framework that improves the classification ability of unlabeled data by combining unsupervised deep clustering and semi-supervised classification together.	Its importance lies in increasing the powerful and classification performance, reduced overfitting caused by bounded labeled data, and efficiently use of large-scale, unlabeled data in medical image analysis.
Mukhlif et al. [23]	Dual Transfer Learning (DTL)	Reduces domain mismatch between source and target domains; fine-tunes with unclassified and classified images.	Limited performance boost without data augmentation; domain-specific knowledge may be required for fine-tuning.	Addressed the problem of inefficient system transfer between pre-trained clinical target models and restricted medical imaging tasks, leading to poor clinical image classification performance.	Its objective is to introduce a new approach that accurately adjusts pre-trained CNNs to unclassified and labeled medical images through dual transfer learning (DTL) to bridge the gap between source and target domain mismatch.	It is important in enhancing the classification performance of breast cancer and skin cancer images, reduces reliance on large-scale classification datasets, and provides a highly efficient domain-specific transfer learning approach.

(Continued)

Table 6. Continued.

Authors	Methods/ Techniques	Strengths	Weaknesses	Problems Addressed	Objectives	Importance
Tharun et al. [106]	Convolutional Neural Networks (CNN)	Outperformed traditional handcrafted feature-based methods; highly accurate for both binary and multiclass classification.	Requires large labeled datasets and high computational resources for training; potential for overfitting with small datasets.	Addressing the problems in breast cancer diagnosis due to reliance on traditional methods for histological diagnosis of the disease by medical experts, errors in manual analysis, and difficulty in accurately classifying tumor types.	The objective is to develop a classification system based on the use of convolutional neural networks (CNNs) for accurate diagnosis of binary (benign vs. malignant) and multiclass (subtypes within benign and malignant categories) breast cancer.	Its importance has emerged in supporting early detection and treatment of breast cancer, as well as reducing the burden on pathologists and providing an objective classification of tumors.
Mukhlif et al. [107]	Transfer Learning with ImageNet	Reduces reliance on ImageNet features by incorporating unclassified images of the same disease, improving model adaptability.	Effectiveness may vary depending on the amount and quality of unclassified medical images available.	Address the problem of the lack of enough labeled medical images to train deep learning models, and reduced reliance on ImageNet features by incorporating unclassified images of the same disease.	Its main purpose is to improve the classification of breast cancer tissue images by integrating unclassified images of the same disease, using pairwise transfer learning technology, and improving data augmentation methods.	Provide a more efficient approach to medical imaging by achieving high classification accuracy while preserving image details without correction-based segmentation.

competitive accuracy. This ensures the ability to operate in resource-constrained environments.

Integrated correction with fine-tuning: While patch-based methods are commonly used (e.g., Nassima et al. [35], Anupama et al. [102]) to reduce computational complexity, optimizing patch extraction and ensuring minimal loss of key features of the image is critical to achieving a balance between accuracy and interpretability.

Combining hand-designed features with deep features: Combining hand-designed features with deep learning-based features (e.g., Amr et al. [105]) enhances the robustness of the model, providing accuracy in classification results while using features known in the field to achieve interpretability.

It becomes possible, then, to address the challenges posed by breast cancer classification while achieving a practical balance between accuracy, interpretability, and computational efficiency by adopting these strategies.

8. Future outlook in breast cancer staging research

Future research is directed towards developing advanced breast cancer staging methods using innovative techniques such as deep learning, artificial intelligence, and unsupervised learning to improve diagnosis accuracy and personalize treatment. The work includes improving classification and feature extraction algorithms and using diverse datasets to increase the effectiveness of predictive models. Attention is focused on developing multi-category taxonomy, creating interpretable models, applying transfer learning, and integrating multimodal data to enhance personalized treatment. In addition, data augmentation techniques such as convolutional neural networks (CNNs) and generative adversarial networks (GANs) are used to overcome the limitations of small data sets. To achieve these goals, it fosters collaboration between researchers and healthcare providers to accelerate innovation and transform breast cancer treatment.

9. Conclusions

Accurate diagnosis and treatment of breast cancer depend on advanced classification of complex histological images. This study highlights the importance of using modern technologies to improve breast cancer staging, and also addresses the challenges associated with analyzing tissue images resulting from breast cancer biopsies. By taking advantage of innovative technology, such as deep learning and transfer learning, and using resources such as the ICIAR 2018 database, researchers have made significant progress toward reliable classification of breast cancer subtypes. These developments contribute to enhancing personalized treatments and improving clinical decision making. Future directions emphasize the importance of integrating histological images with genomic and clinical data in order to achieve a comprehensive understanding of tumor biology. Integrating multimodal data, as well as data augmentation techniques and collaborative platforms for information sharing, is key to overcoming the limitations of datasets and enhancing the effectiveness of models. Moreover, the development of lightweight and efficient models contributes to ensuring their accessibility in resource-limited environments, while unsupervised learning shows promising potential in discovering new patterns without relying on labeled datasets. These advances pave the way for more accurate diagnoses, improved patient outcomes, and the development of innovative treatment models. This survey is a comprehensive resource that reviews classification methods, identifies challenges, and suggests future research paths. It also highlights the importance of cooperation between researchers and healthcare providers in promoting technological progress that contributes to improving the diagnosis and staging of breast cancer.

Funding

None

Acknowledgement

None

Conflicts of interest

The author declares no conflict of interest.

References

1. H. J. Mohammed, A. A. Nafea, H. K. Almulla, S. A. S. Aliesawi, and M. M. Al-Ani, “An effective hybrid model for skin cancer detection using transfer learning,” in *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 840–845, 2023.
2. G. Sajiv, G. Ramkumar, S. Shanthi, A. Chinnathambi, and S. A. Alharbi, “Predicting breast cancer risk from histopathology images using hybrid deep learning classifier,” *Med. Eng. Phys.*, p. 104149, 2024.
3. A. R. Hmeed, S. A. Aliesawi, and W. M. Jasim, “Enhancement of the U-net architecture for MRI brain tumor segmentation,” in *Next Generation of Internet of Things: Proceedings of ICNGIoT 2021*, pp. 353–367, 2021.
4. O. J. Kadhim, A. A. Nafea, S. A. S. Aliesawi, and M. M. Al-Ani, “Ensemble model for prostate cancer detection using MRI images,” in *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 492–497, 2023.
5. K. Kumari, P. Groza, and F. Aguilo, “Regulatory roles of RNA modifications in breast cancer,” *NAR cancer*, vol. 3, no. 3, p. zcab036, 2021.
6. M. I. Meyer, A. Galdran, P. Costa, A. M. Mendonça, and A. Campilho, “Deep convolutional artery/vein classification of retinal vessels,” in *Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15*, 2018, pp. 622–630.
7. A. P. Vekariya, “A deep learning based pipeline for metastatic breast cancer classification from whole slide images (WSI).” 2023.
8. B. Abhisheka, S. K. Biswas, and B. Purkayastha, “A comprehensive review on breast cancer detection, classification and segmentation using deep learning,” *Arch. Comput. Methods Eng.*, vol. 30, no. 8, pp. 5023–5052, 2023.
9. Y. Xie *et al.*, “Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives,” *Diagnostics*, vol. 12, no. 8, p. 1850, 2022.
10. A. Williams, A. Cunningham, H. Hutchings, D. A. Harris, M. D. Evans, and D. Harji, “Quality of internet information to aid patient decision making in locally advanced and recurrent rectal cancer,” *Surg.*, vol. 20, no. 6, pp. e382–e391, 2022.
11. P. Basu, L. Zhang, R. Hariprasad, A. L. Carvalho, and A. Barchuk, “A pragmatic approach to tackle the rising burden of breast cancer through prevention & early detection in countries ‘in transition,” *Indian J. Med. Res.*, vol. 152, no. 4, pp. 343–355, 2020.
12. A. Ashurov, S. A. Chelloug, A. Tselykh, M. S. A. Muthanna, A. Muthanna, and M. S. A. M. Al-Gaashani, “Improved breast cancer classification through combining transfer learning and attention mechanism,” *Life*, vol. 13, no. 9, p. 1945, 2023.
13. M. M. Srikantamurthy, V. P. S. Rallabandi, D. B. Dudekula, S. Natarajan, and J. Park, “Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning,” *BMC Med. Imaging*, vol. 23, no. 1, p. 19, 2023.
14. X. Dai *et al.*, “Breast cancer intrinsic subtype classification, clinical use and future trends,” *Am. J. Cancer Res.*, vol. 5, no. 10, p. 2929, 2015.
15. A. Yaqoob, R. Musheer Aziz, and N. K. verma, “Applications and techniques of machine learning in cancer classification: A systematic review,” *Human-Centric Intell. Syst.*, vol. 3, no. 4, pp. 588–615, 2023.

16. W. R. Drioua, N. Benamrane, and L. Sais, "Breast cancer histopathological images segmentation using deep learning," *Sensors*, vol. 23, no. 17, p. 7318, 2023.
17. R. Rashmi, K. Prasad, and C. B. K. Udupa, "Breast histopathological image analysis using image processing techniques for diagnostic purposes: A methodological review," *J. Med. Syst.*, vol. 46, no. 1, p. 7, 2022.
18. A. A. A. Lateef, A. S. Abdalkafor, and A. A. Nafea, "Optimized KNN algorithm for diabetic retinopathy classification with PCA-based data fusion and cuckoo search optimization."
19. M. K. Santos, J. R. Ferreira, D. T. Wada, A. P. M. Tenório, M. H. Nogueira-Barbosa, and P. M. de A. Marques, "Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: Advances in imaging towards to precision medicine," *Radiol. Bras.*, vol. 52, no. 06, pp. 387–396, 2019.
20. S. Y. Mohammed, "Enhancing COVID-19 patients detection using deep transfer learning technique through x-ray chest images," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 32, no. 1, pp. 290–302, 2023.
21. S. Xu *et al.*, "Integrative analysis of histopathological images and chromatin accessibility data for estrogen receptor-positive breast cancer," *BMC Med. Genomics*, vol. 13, pp. 1–12, 2020.
22. B. Xiao and C. Lu, "Semi-supervised medical image classification combined with unsupervised deep clustering," *Appl. Sci.*, vol. 13, no. 9, p. 5520, 2023.
23. A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed, "Incorporating a novel dual transfer learning approach for medical images," *Sensors*, vol. 23, no. 2, p. 570, 2023.
24. A. A. Nafea, A.-M. Manar, K. M. A. Alheeti, M. S. I. Alsumaidaie, and M. M. AL-Ani, "A hybrid method of 1D-CNN and machine learning algorithms for breast cancer detection," *Baghdad Sci. J.*, 2024.
25. S. Łukasiewicz, M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, and A. Stanisławek, "Breast cancer—epidemiology, risk factors, classification, prognostic markers, and current treatment strategies—an updated review," *Cancers (Basel).*, vol. 13, no. 17, p. 4287, 2021.
26. A. A. Nafea, A.-M. Manar, M. M. AL-Ani, and N. Omar, "A review on adverse drug reaction detection techniques," *ARO-THE Sci. J. KOYA Univ.*, vol. 12, no. 1, pp. 143–153, 2024.
27. J. Makki, "Diversity of breast carcinoma: Histological subtypes and clinical relevance," *Clin. Med. insights Pathol.*, vol. 8, p. CPath-S31563, 2015.
28. M. Nasser and U. K. Yusof, "Deep learning based methods for breast cancer diagnosis: A systematic review and future direction," *Diagnostics*, vol. 13, no. 1, p. 161, 2023.
29. J. Xie, R. Liu, J. Luttrell IV, and C. Zhang, "Deep learning based analysis of histopathological images of breast cancer," *Front. Genet.*, vol. 10, p. 80, 2019.
30. D. Bardou, K. Zhang, and S. M. Ahmad, "Classification of breast cancer based on histology images using convolutional neural networks," *Ieee Access*, vol. 6, pp. 24680–24693, 2018.
31. F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 2560–2567.
32. Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Sci. Rep.*, vol. 7, no. 1, p. 4172, 2017.
33. S. Rezaei *et al.*, "Gland segmentation in histopathology images using deep networks and handcrafted features," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 1031–1034.
34. R. Yan *et al.*, "A hybrid convolutional and recurrent deep neural network for breast cancer pathological image classification," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, pp. 957–962.
35. N. Dif and Z. Elberrichi, "A new intra fine-tuning method between histopathological datasets in deep learning," *Int. J. Serv. Sci. Manag. Eng. Technol.*, vol. 11, no. 2, pp. 16–40, 2020.
36. T. Kausar, Y. Lu, and A. Kausar, "Breast cancer diagnosis using lightweight deep convolution neural network model," *IEEE Access*, vol. 11, pp. 124869–124886, 2023.
37. R. Yan *et al.*, "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52–60, 2020.
38. M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.
39. X. Li, M. Radulovic, K. Kanjer, and K. N. Plataniotis, "Discriminative pattern mining for breast cancer histopathology image classification via fully convolutional autoencoder," *IEEE Access*, vol. 7, pp. 36433–36445, 2019.
40. S. Imran, B. A. Lodhi, and A. Alzahrani, "Unsupervised method to localize masses in mammograms," *IEEE Access*, vol. 9, pp. 99327–99338, 2021.
41. S. Lee, C. Farley, S. Shim, W.-S. Yoo, Y. Zhao, and W. Choi, "Unsupervised learning of deep-learned features from breast cancer images," in *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2020, pp. 740–745.
42. A. Tiard *et al.*, "Stain-invariant self supervised learning for histopathology image analysis," *arXiv Prepr. arXiv2211.07590*, 2022.
43. Y. Zhao, J. Zhang, D. Hu, H. Qu, Y. Tian, and X. Cui, "Application of deep learning in histopathology images of breast cancer: A review," *Micromachines*, vol. 13, no. 12, p. 2197, 2022.
44. F. Abdullakutty, Y. Akbari, S. Al-Maadeed, A. Bouridane, and R. Hamoudi, "Advancing histopathology-based breast cancer diagnosis: Insights into multi-modality and explainability," *arXiv Prepr. arXiv2406.12897*, 2024.
45. A. Modi *et al.*, "Multi-stain multi-level convolutional network for multi-tissue breast cancer image segmentation," *arXiv Prepr. arXiv2406.05828*, 2024.
46. Y. Huang and A. C. Chung, "Improving high resolution histology image classification with deep spatial fusion network," in *Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY2018, and 5th International Workshop, OMIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018, Proceedings 5*, 2018, pp. 19–26.
47. D. J. Ho *et al.*, "Deep multi-magnification networks for multi-class breast cancer image segmentation," *Comput. Med. Imaging Graph.*, vol. 88, p. 101866, 2021.
48. B. Wei, Z. Han, X. He, and Y. Yin, "Deep learning model based breast cancer histopathological image classification," in *2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2017, pp. 348–353.
49. G. Rea *et al.*, "Beyond visual interpretation: Quantitative analysis and artificial intelligence in interstitial lung disease diagnosis 'expanding horizons in radiology," *Diagnostics*, vol. 13, no. 14, p. 2333, 2023.
50. J. Vizcarra, R. Place, L. Tong, D. Gutman, and M. D. Wang, "Fusion in breast cancer histology classification," in

Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, 2019, pp. 485–493.
51. A. Bagchi, P. Pramanik, and R. Sarkar, “A multi-stage approach to breast cancer classification using histopathology images,” *Diagnostics*, vol. 13, no. 1, p. 126, 2022.
52. A. A. Mukhlif, B. Al-Khateeb, and M. Mohammed, “Classification of breast cancer images using new transfer learning techniques,” *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 1, pp. 167–180, 2023.
53. M. A. Wakili *et al.*, “Classification of breast cancer histopathological images using DenseNet and transfer learning,” *Comput. Intell. Neurosci.*, vol. 2022, 2022.
54. E. Pfaehler *et al.*, “PET segmentation of bulky tumors: Strategies and workflows to improve inter-observer variability,” *PLoS One*, vol. 15, no. 3, p. e0230901, 2020.
55. C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, “Key challenges for delivering clinical impact with artificial intelligence,” *BMC Med.*, vol. 17, pp. 1–9, 2019.
56. J. S. Ahn *et al.*, “Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine,” *J. Breast Cancer*, vol. 26, no. 5, p. 405, 2023.
57. A. Pulumati, A. Pulumati, B. S. Dwarakanath, A. Verma, and R. V. L. Papineni, “Technological advancements in cancer diagnostics: Improvements and limitations,” *Cancer Rep.*, vol. 6, no. 2, p. e1764, 2023.
58. E. Alizadeh, J. Castle, A. Quirk, C. D. L. Taylor, W. Xu, and A. Prasad, “Cellular morphological features are predictive markers of cancer cell state,” *Comput. Biol. Med.*, vol. 126, p. 104044, 2020.
59. S. Di Cataldo and E. Ficarra, “Mining textural knowledge in biological images: Applications, methods and trends,” *Comput. Struct. Biotechnol. J.*, vol. 15, pp. 56–67, 2017.
60. R. D. Chitalia and D. Kontos, “Role of texture analysis in breast MRI as a cancer biomarker: A review,” *J. Magn. Reson. Imaging*, vol. 49, no. 4, pp. 927–938, 2019.
61. Z. Jaadi, “A step-by-step explanation of principal component analysis (PCA),” *Retrieved June*, vol. 7, no. 2021, p. 470, 2021.
62. A. Tripathi, “A complete guide to principal component analysis—PCA in machine learning,” *Mediu. Towar. Data Sci.*, vol. 11, 2019.
63. I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene selection for cancer classification using support vector machines,” *Mach. Learn.*, vol. 46, pp. 389–422, 2002.
64. X. Chen and J. C. Jeong, “Enhanced recursive feature elimination,” in *Sixth International Conference on Machine Learning and Applications (ICMLA2007)*, 2007, pp. 429–435.
65. S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: Past, present, and future,” *Multimed. Tools Appl.*, vol. 80, pp. 8091–8126, 2021.
66. G. Park, K.-N. Hong, and H. Yoon, “Vision-based structural FE model updating using genetic algorithm,” *Appl. Sci.*, vol. 11, no. 4, p. 1622, 2021.
67. N. Pudjihartono, T. Fadason, A. W. Kempa-Liehr, and J. M. O’Sullivan, “A review of feature selection methods for machine learning-based disease risk prediction,” *Front. Bioinforma.*, vol. 2, p. 927312, 2022.
68. S. Perumal and T. Velmurugan, “Preprocessing by contrast enhancement techniques for medical images,” *Int. J. Pure Appl. Math.*, vol. 118, no. 18, pp. 3681–3688, 2018.
69. O. S. Al-Kadi, “A gabor filter texture analysis approach for histopathological brain tumor subtype discrimination,” *arXiv Prepr. arXiv1704.05122*, 2017.
70. Y. Amethiya, P. Pipariya, S. Patel, and M. Shah, “Comparative analysis of breast cancer detection using machine learning and biosensors,” *Intell. Med.*, vol. 2, no. 2, pp. 69–81, 2022.
71. W. Dean, “Emerging advances to transform histopathology using virtual staining,” *BME Front.*, 2020.
72. T. Araújo *et al.*, “Classification of breast cancer histology images using convolutional neural networks,” *PLoS One*, vol. 12, no. 6, p. e0177544, 2017.
73. G. Dileep and S. G. G. Gyani, “Artificial intelligence in breast cancer screening and diagnosis,” *Cureus*, vol. 14, no. 10, 2022.
74. R. A. Dar, M. Rasool, and A. Assad, “Breast cancer detection using deep learning: Datasets, methods, and challenges ahead,” *Comput. Biol. Med.*, vol. 149, p. 106073, 2022.
75. N. Fico *et al.*, “Breast imaging physics in mammography (Part I),” *Diagnostics*, vol. 13, no. 20, p. 3227, 2023.
76. M. Veta, J. P. W. Pluim, P. J. Van Diest, and M. A. Viergever, “Breast cancer histopathology image analysis: A review,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1400–1411, 2014.
77. A.-A. Nahid, M. A. Mehrabi, and Y. Kong, “Histopathological breast cancer image classification by deep neural network techniques guided by local clustering,” *Biomed Res. Int.*, vol. 2018, 2018.
78. M. Lee, “Deep learning techniques with genomic data in cancer prognosis: A comprehensive review of the 2021–2023 literature,” *Biology (Basel).*, vol. 12, no. 7, p. 893, 2023.
79. C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *J. big data*, vol. 6, no. 1, pp. 1–48, 2019.
80. S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2009.
81. A. Das, M. S. Nair, and D. S. Peter, “Batch mode active learning on the Riemannian manifold for automated scoring of nuclear pleomorphism in breast cancer,” *Artif. Intell. Med.*, vol. 103, p. 101805, 2020.
82. A. Vahadane *et al.*, “Structure-preserving color normalization and sparse stain separation for histological images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 8, pp. 1962–1971, 2016.
83. A. Madabhushi and G. Lee, “Image analysis and machine learning in digital pathology: Challenges and opportunities,” *Med. Image Anal.*, vol. 33, pp. 170–175, 2016.
84. E. Çallı, E. Sogancioglu, B. van Ginneken, K. G. van Leeuwen, and K. Murphy, “Deep learning for chest X-ray analysis: A survey,” *Med. Image Anal.*, vol. 72, p. 102125, 2021.
85. V. C. Nitesh, “SMOTE: Synthetic minority over-sampling technique,” *J Artif Intell Res*, vol. 16, no. 1, p. 321, 2002.
86. Z.-H. Zhou and X.-Y. Liu, “Training cost-sensitive neural networks with methods addressing the class imbalance problem,” *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 1, pp. 63–77, 2005.
87. H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.
88. A. G. Howard, “Mobilenets: Efficient convolutional neural networks for mobile vision applications,” *arXiv Prepr. arXiv1704.04861*, 2017.
89. J. Chen and X. Ran, “Deep learning with edge computing: A review,” *Proc. IEEE*, vol. 107, no. 8, pp. 1655–1674, 2019.
90. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Artificial Intelligence and Statistics*, 2017, pp. 1273–1282.

91. G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
92. M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.
93. A. Dosovitskiy, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative unsupervised feature learning with convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.
94. C. P. Nguyen, A. H. Vo, and B. T. Nguyen, "Breast cancer histology image classification using deep learning," in *2019 19th International Symposium on Communications and Information Technologies (ISCIT)*, 2019, pp. 366–370.
95. T. A. Toma *et al.*, "Breast cancer detection based on simplified deep learning technique with histopathological image using BreaKHis database," *Radio Sci.*, vol. 58, no. 11, pp. 1–18, 2023.
96. R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights Imaging*, vol. 9, pp. 611–629, 2018.
97. S. Zakareya, H. Izadkhah, and J. Karimpour, "A new deep-learning-based model for breast cancer diagnosis from medical images," *Diagnostics*, vol. 13, no. 11, p. 1944, 2023.
98. L. Jose, S. Liu, C. Russo, A. Nadort, and A. Di Ieva, "Generative adversarial networks in digital pathology and histopathological image processing: A review," *J. Pathol. Inform.*, vol. 12, no. 1, p. 43, 2021.
99. I. A. Cree *et al.*, "The international collaboration for cancer classification and research," *Int. J. Cancer*, vol. 148, no. 3, pp. 560–571, 2021.
100. H. Le Minh, M. M. Van, T. T. Dinh, T. T. Dac, and T. Van Lang, "Automatic diagnosis of breast cancer in histology images using deep convolutional neural networks."
101. Y. Guo, H. Dong, F. Song, C. Zhu, and J. Liu, "Breast cancer histology image classification based on deep neural networks," in *Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15*, 2018, pp. 827–836.
102. M. A. Anupama, V. Sowmya, and K. P. Soman, "Breast cancer classification using capsule network with preprocessed histology images," in *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 143–147.
103. S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Breast cancer diagnosis with transfer learning and global pooling," in *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, 2019, pp. 519–524.
104. R. P. da Silva Neto and A. O. de Carvalho Filho, "Automatic classification of breast lesions using transfer learning," *IEEE Lat. Am. Trans.*, vol. 17, no. 12, pp. 1964–1969, 2019.
105. A. H. Hassan, M. E. Wahed, M. A. Atiea, and M. S. Metwally, "A hybrid approach for classification breast cancer histopathology images," *Front. Sci. Res. Technol.*, vol. 3, no. 1, pp. 1–10, 2022.
106. T. Kumar, M. Chandran, G. S. Priyatharsini, S. Geetha, and J. Jayaprakas, "Breast cancer detection using machine learning classifier," *Cryobiology*, vol. 114, p. 104820, 2024.
107. A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed, "Breast cancer images classification using a new transfer learning technique," *Iraqi J Comput Sci Math*, vol. 4, no. 1, pp. 167–180, 2023.