Alexandria Engineering Journal 89 (2024) 210–223

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

ORIGINAL ARTICLE

Machine learning, IoT and 5G technologies for breast cancer studies: A review

Havva Elif Saroğlu [a], Ibraheem Shayea [a,*], Bilal Saoud [b,*], Marwan Hadri Azmi [c,*], Ayman A. El-Saleh [d], Sawsan Ali Saad [e], Mohammad Alnakhli [f]

[a] Electronics & Communications Engineering Department, Faculty of Electrical and Electronics Engineering, Istanbul Technical University (ITU), 34469, Istanbul, Turkey
[b] Electrical Engineering Department, Sciences and Applied Sciences Faculty, Bouira University, 10000, Bouira, Algeria
[c] Wireless Communication Centre, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia
[d] Department of Electronics and Communication Engineering, College of Engineering, A'Sharqiyah University, Ibra 400, Oman
[e] Department of Computer Engineering, University of Ha'il, Ha'il, 55211, Saudi Arabia
[f] Electrical Engineering Dept., College of Engineering, Prince Sattam Bin Abdulaziz University, Wadi Addwasir, 11991, Saudi Arabia

A R T I C L E I N F O

Keywords:
Breast cancer
Classification algorithms
Deep learning
Histopathological images
Machine learning
Mammogram
IoT
5G

A B S T R A C T

Cancer is a life-threatening ailment characterized by the uncontrolled proliferation of cells. Breast cancer (BC) represents the most highly infiltrative neoplasms and constitutes the primary cause of mortality in the female population due to cancer-related complications. Consequently, the imperative for early detection and prognosis has emerged as a means to enhance long-term survival rates and mitigate mortality. Emerging artificial intelligence (AI) technologies are being utilized to aid radiologists in the analysis of medical images, resulting in enhanced outcomes for individuals diagnosed with cancer. The purpose of this survey is to examine peer-reviewed computer-aided diagnosis (CAD) systems that have been recently developed and utilize machine learning (ML) and deep learning (DL) techniques for the diagnosis of BC. The survey aims to compare these newly developed systems with previously established methods and provide technical details, as well as the advantages and disadvantages associated with each model. In addition, this paper addresses several unresolved matters, areas of research that require further exploration, and potential avenues for future investigation in the realm of advanced computer-aided design (CAD) models utilized in the interpretation of medical images. Furthermore, the integration of Internet of Things (IoT) in BC research and treatment holds immense significance by facilitating real-time monitoring and personalized healthcare solutions. IoT devices, such as wearable sensors and smart implants, enable continuous data collection, empowering healthcare professionals to track patients' vital signs, response to treatment, and overall health trends, fostering more proactive and tailored approaches to BC management. Moreover, the advent of 5G technology in BC applications promises to revolutionize communication speeds and data transfer, enabling rapid and seamless transmission of large medical datasets. This high-speed connectivity enhances the efficiency of remote diagnostics, telemedicine, and collaborative research efforts, ultimately accelerating the pace of innovation and improving patient outcomes in BC care. The present study aims to examine various classifiers utilized in ML and DL methodologies for the purpose of diagnosing BC. Research findings have demonstrated that DL has superior performance compared to standard ML methods in the context of BC diagnosis, particularly when the dataset is extensive. The existing body of research indicates that there are significant gaps in knowledge that need to be addressed in order to enhance healthcare outcomes in the future. These gaps highlight the pressing need for both practical and scientific research in the field. Finally, IoT and 5G will be how they can be used in order to enhance BC detection, treatment and patient care.

* Corresponding authors.
E-mail addresses: shayea@itu.edu.tr (I. Shayea), bilal340@gmail.com (B. Saoud), hadri@utm.my (M.H. Azmi), ayman.elsaleh@asu.edu.om (A.A. El-Saleh), sa.saad@uoh.edu.sa (S.A. Saad), m.alnakhli@psau.edu.sa (M. Alnakhli).

https://doi.org/10.1016/j.aej.2024.01.043
Received 6 October 2023; Received in revised form 22 December 2023; Accepted 13 January 2024
Available online 25 January 2024

1. Introduction

Cancer is one of the most common and deadly disease around the world and it has numerous different types. Recently, studies proved that Breast Cancer (BC) is the most prevalent cancer that is seen in women around the world with an estimated number of 2.3 million new cases and 685.000 deaths related to breast cancer [1]. Also, in United States (US) statistics have shown that more than a quarter million people were diagnosed with BC with a fatality number of 41.690 deaths, making BC the most frequent and fatal cancer type [2]. Furthermore, breast cancer accounts for approximately a quarter of all cancer diagnoses in women. [3]. Based on these statistics, we can clearly understand that BC is a big issue in public health and needs to be taken into consideration to prevent further casualties. However, these concerns have led to advancements in the use of modern technologies such as machine learning (ML) and deep learning for BC diagnosis.

Different types of BC require different therapies. While surgery, radiotherapy and chemotherapy are some of the most common methods used in the treatment of breast cancer patients, there are also modern treatment alternatives such as immunotherapy or targeted therapy, the main approach on reducing the mortality rates of BC is early diagnosis. It is well known that for patient diagnosed with BC in its early stages has around 5-year survival rate of about 99% [4]. And what is meant by early diagnosis is detecting malignant cells before they spread to other parts of the body such as lymph nodes, lungs or other organs. Physical examination, ultrasound, X-ray Mammography, Magnetic Resonance Imaging (MRI), lab results or biopsy are utilized for early detection of breast cancer. Histopathological analysis can usually determine the exact type of cancer; hence it is considered the golden standard for BC diagnosis. However, it is a lengthy procedure that necessitates highly qualified and experienced pathologists. Additionally, interpreting the resulting images from the MRI, ultrasound or mammography can be quite challenging for the radiology specialists. Both of these methods for detecting and classifying breast cancer rely heavily on human mind, making them extremely error-prone.

Studies based on ML and deep learning approaches to breast cancer diagnosis are increasing in order to diagnose and classify the disease as quickly as possible in order to start therapies before the spreading of the disease. There are various definitions to explain the exact process of ML. It is a sub layer of Artificial Intelligence (AI) that may be characterized as a group of algorithms that teach computers how to detect patterns in huge, complex data sets using statistical and probabilistic models. ML, according to IBM, is based on mimicking human learning on computers step by step using various algorithms and massive datasets [5].

ML techniques are classified as supervised, unsupervised, semi-supervised and reinforcement learning based on whether or not the data is tagged. Deep learning, on the other hand, is a subset of ML that focuses on deep neural network-based learning methods. Deep neural networks (DNN) are models which imitate the learning procedure of human brain. In recent years, research into the use of these algorithms in healthcare, particularly in cancer researches has surged due to their pattern-detection capabilities. These algorithms can interpret medical images, pathology results, and create predictive analytics by detecting hidden patterns in large data sets or categorizing images based on their learning, making them ideal tools for BC diagnosis. The major goal of combining technology with medical diagnosis and treatment here is to determine the type and prognosis of BC for each patient with the least amount of errors conceivable in order to create the optimal treatment plan and avoiding high fatality rates. In order to gain a better knowledge about BC, some researchers are eager to employ a variety of ML techniques to predict recurrence and 5-year survival rates.

Physicians have been showing gradual progress in cancer studies over the past decades, still the definite prognosis of BC patients remains as a demanding process [6]. This challenge has resulted in advancements in the ML studies for the classification and the prediction of breast cancer. Since, ML studies in breast cancer are a relatively new research field, there are many unknown factors and unanswered issues concerning the existing models' performance. Given how important it is for a cancer patient to be diagnosed with smallest possible error for the most accurate treatment plan, so that the person of interest has the best chance of survival. It is crucial to select the suitable model when using ML algorithms [7]. However, studies have demonstrated that models for BC based on ML still require adequate validation before they can be utilized in clinical care on a regular basis [6].

There are numerous studies related to ML for BC in the current literature [7–11], [13], [19], [21,22], [25,30], [32–34], [52–54]. One of the most prevalent studies in this field is using ML algorithms to classify medical images as benign or malignant. The authors of [8] utilized Convolutional Neural Networks (CNN) with 1 input layer, 28 hidden layers and one output layer to categorize the mammographic images in a data set obtained from Mammographic Image Analysis Society (MIAS). The system was called the Convolutional Neural Network Improvement for Breast Cancer Classification (CNNI-BCC) and it achieved a sensitivity and accuracy of approximately 90% when classifying the images as benign, malignant or normal [8]. Although, mammography is considered to be the golden standard for BC diagnosis. To decrease the false positive and false negative rates in analysis of ultrasound images, authors in [9] utilized Artificial Neural Network Classifiers to classify 184 breast sonograms as benign or malignant.

In [10], scientists offered a systematic evaluation of current research in thermographic imaging and how the data might be interpreted using deep learning approaches such as CNN, in contrast to standard screening methods in BC diagnosis. Other than classifying the screening results, there are also several studies in categorizing the histopathological images [11], [21], [31–33]. The authors in [11] introduced a highly efficient transfer learning framework to diagnose the specific molecular sub-type of sample tissues with multiclass classification and to determine whether a tissue sample is malignant or not with binary classification. The model worked perfectly for the binary classification technique and displayed 98% accuracy for detecting the exact sub-type. Other than the categorization of the cancer type, estimating the recurrence and 5-years survival rates of patients is crucial for BC studies. In [12], it is stated that ML methods can increase the estimation of relapse in cancer patients up to 25% compared to conventional techniques. ML models based on random forest algorithm were able to detect BC recurrence by making use of clinicopathological features in 3 months advance. While there are numerous studies on ML in BC in the current literature, it is still critical to create models that are applicable to routine health care and valid all over the globe.

This paper provides a review of the most commonly used methods and algorithms for BC classification. The basic concepts of ML in BC will be identified and explored in this paper. Since ML studies are being investigating intensively, particularly in cancer studies, it is necessary to summarize previous works and commonly used techniques in order to improve the research in this field. Ultimately, this paper will assist future researchers in understanding the steps necessary to begin a project in this field and will emphasize the most important aspects of currently used methods.

This paper is organized as follows. Section 2 outlines how to classify breast cancer from a medical perspective. Algorithms and methodologies generally utilized in ML studies for BC are covered in depth in Section 3. Section 4 contains details on current and previous ML studies for BC screening, pathology and predictive analysis. The future trends in this field are discussed in Section 5. Finally, section 6 concludes the paper.

2. BC classification

For every patient, determining the type of cancer they have is critical for treatment and prognosis. Accurately classifying the cancer's variation enables physicians to come up with the optimal treatment plan and predict the survival chances of the patient. Fig. 1 shows the procedures

Fig. 1. BC diagnosis methods.

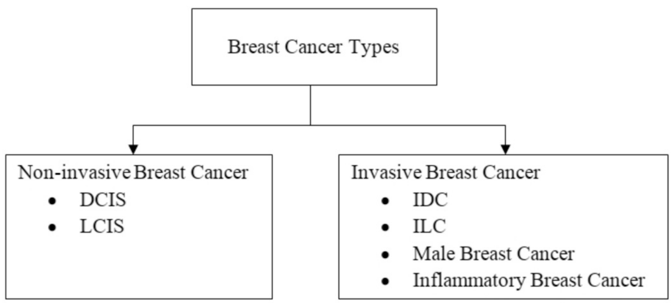

Fig. 2. BC types divided by two main category non-invasive and invasive cancers.

for breast cancer diagnosis. There are several aspects that can be considered while determining the style of the cancer, but the most common are the type and the stage [1,5].

2.1. Based on type

BC types can usually be split into two main categories: non-invasive and invasive BC. There is also some sub-categorization based on status of the hormone receptors, proteins or the genes of the cancer cells which affects the treatment plan. Fig. 2 displays the main types of BC [37].

2.1.1. Non-invasive BC

Non-invasive BC, in other words carcinoma in situ or pre-cancers, means that the malignant cells are obtained inside their generation location. The first one is Ductal Carcinoma in Situ (DCIS). It is used to describe the situation when cancer cells are contained inside the milk ducts and the normal breast tissue remains untouched and healthy [13]. The other one is Lobular Carcinoma in Situ (LCIS), which can be understood from its name. It is when the cancer is in the lobule. Both of these cancer types are mostly seen as pre-cancers which are not dangerous but as warning signs that may develop into invasive cancers.

2.1.2. Invasive BC

Invasive cancers are when the cancer does not stay enclosed at the same place and invades the healthy breast tissue [13,37]. Most of breast cancers seen in women are invasive cancers. Invasive Ductal Carcinoma (IDC) is the most frequently diagnosed cancer among BC types which makes up almost 80% of the BC diagnosis. For IDC patients, the cancer originates in the milk tubes and spreads to the breast or other organs. The second most common type is Infiltrating Lobular Carcinoma (ILC). Cancer starts in the lobules or the lobes and spreads the body. ILC usually responds well to hormone therapy. There are other types of BC which are seen less. Male breast cancer is incredibly rare and it is always ductal carcinoma. Inflammatory Breast Cancer is also a very unique and aggressive type, which mostly causes breast to swallow and become red. There are also Recurrent and Metastatic Breast Cancers meaning that the cancer came back after the treatment or it has spread the other organs like lungs, bones and brain.

2.1.3. Molecular sub-types

Categorization of cancer cells based on their specific genes or receptors called molecular sub-types. This process is necessary to understand the behavior of tumor and determine the treatment procedure. There are 4 molecular sub-types classified based on cells' status on progesterone-receptor (PR), estrogen-receptor (ER) and human epidermal growth factor receptor 2 (HER2). These are Luminal A, Luminal B, HER-2 enriched and triple-negative [14].

Luminal A is ER and PR positive and HER2 negative. These cells grow with the existence of estrogen and progesterone. Luminal A tumor cells spread and grow smaller than the other sub-types. It is also the most prevalent sub-type that is seen in BC patients. Lowering the amount of estrogen and progesterone are usually included to the treatment plan.

Unlike Luminal A, Luminal B is triple positive. It is ER+, PR+ and HER2+. This results in faster growth and poorer diagnosis. Usually 30% of BC patients are diagnosed as Luminal B making it the second most common sub-type.

HER2 enriched suggested for the situation when the cells are ER and PR negative, but HER2 positive. The tumors tend to grow faster. But it generally responds to HER2 targeted therapy. It has poor prognosis. Triple-negative as it can be clearly understood from its name suggests that the cells are negative for ER, PR and HER2. Treatment includes chemotherapy. It is considered as an aggressive type.

2.2. Based on stage

While sorting the class of the breast cancers they are classified into 5 stages from 0 to 4 [15]. The stage of the cancer gives an understanding about the prognosis of the disease. There are lots of markers and specifications to determine the exact stage of the cancer. But to explain briefly, Stage 0 is when the cancer is inside the milk ducts or the glands and has not spread to anywhere else. It means that the cancer is diagnosed at early stages. Stage I suggests invasive cancers meaning that there are malignant cells in the breast tissue. For Stage II, tumor may be growing or spreading to close lymph nodes. Stage III can be described as an advance phase. It is generally used for situations when cancer has not spread to any organs yet but it is in the lymph nodes or the chest wall. And Stage IV is seen as the final stage, it is considered for situations when there are tumors in organs far away from the breast like brain, bones, liver or lungs.

There is also grading assessment to get a grasp of the behavior of the cancer cells. It is done by examining the behavior of the malign cells. Grade 1 is for cells similar to breast tissue and very slowly growing. And Grade 3 is when cells look very different from the normal breast cells and inclined to grow and spread incredibly fast. Grade 2 describes some stage between the Grade 1 and 3.

3. Common algorithms of ML for BC

ML algorithms are frequently used in medical researches to provide prediction diagnosis of numerous diseases. ML approaches are mainly divided into two sub-sets as supervised and unsupervised learning based on their data type, and classification falls under the category of supervised learning. Supervised learning means that the training dataset includes labeled data. To give a brief example, assuming that there is a data set that contains mammograms, if each mammogram is labeled as normal or suspicious then it can be declared that the algorithms which utilize this data base are supervised algorithms. Fig. 3 demonstrates the process flow for BC classifiers which uses histopathological analysis. In the following some ML algorithms will be presented.

3.1. K-nearest neighbors (KNN)

KNN is a slow algorithm which provides high accuracy. It works with similarities between the samples. It assumes that similar models

Fig. 3. Example of classification process for histopathological images.

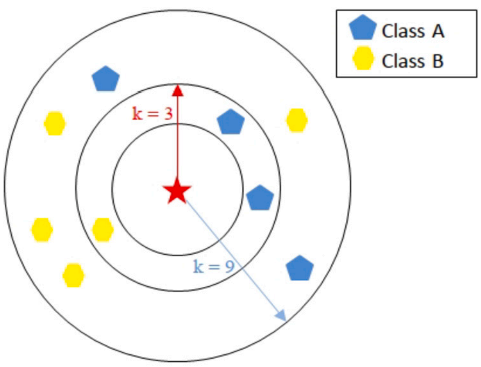

Fig. 4. Graphical form of KNN method.

are close to each other. K is a variable which states the number of samples that the algorithm will take into consideration. Algorithm calculates the distances between the training point and test samples, and then looks at the labels of the K closest sample. The algorithm works on the principle of "majority voting", meaning that it decides the category of the test sample based on the leading class label among the k-nearest neighbors [16,56]. Fig. 4 illustrates the working principle of KNN. The red star in the middle is the test sample, and the algorithm is attempting to classify it by looking at its neighbors. When $k = 3$, the star sample is sorted as a pentagon, because there are two pentagon samples and just one hexagon. Still, assuming that $k = 9$, then number of hexagons will be higher than the pentagon so the sample will be classified as a hexagon. The simplified method is to set $k = 1$, which results in the test sample to take the same label with the closest neighbor. However, this is not generally preferred because it does not go well with big, noisy data sets.

KNN is a simple and effective algorithm used for classification in the context of BC classification. KNN can be used to classify a patient's breast tissue as either malignant or benign based on the features of the tissue. Basically KNN is used based on these steps:

- Gather a dataset of breast tissue samples that have already been classified as malignant or benign.
- Extract features from each sample, such as the size, shape, and texture of the cells.
- Split the dataset into a training set and a test set.
- Choose a value for K, which is the number of nearest neighbors to consider when making a classification.
- For each sample in the test set, calculate the distance between that sample and all of the samples in the training set.
- Select the K samples from the training set that are closest to the test sample.
- Classify the test sample as malignant or benign based on the majority class of its K nearest neighbors.
- Evaluate the accuracy of the algorithm by comparing the predicted classifications to the actual classifications in the test set.

One potential issue with using KNN for BC classification is that it requires a large dataset with many features to achieve good accuracy. Additionally, it may not be as effective at identifying subtle differences between malignant and benign tissue as more advanced algorithms. However, it can be a useful starting point for researchers and clinicians who are interested in using ML for BC classification.

3.2. Naïve Bayes algorithm

Naïve Bayes is named after the mathematician Bayes and it is based on the Bayes Theorem of probability [57,59]. Which is:

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right)P(A)}{P(B)} \tag{1}$$

It works under the assumption that defining properties that will be used in the categorization process are unrelated to one another.

BC classification based on Naïve Bayes algorithm involves using a probabilistic model to predict whether a given BC case is malignant or benign. Naïve Bayes algorithm assumes that the features are independent of each other, which may not be true in reality, but still it is often effective in practice. The algorithm calculates the probability of a particular feature occurring given the class of the BC and multiplies it with the prior probability of the class. This is done for all features and classes, and the class with the highest probability is chosen as the predicted class.

Several studies have applied Naïve Bayes algorithm for BC classification. For example, authors of [57] compared the performance of Naïve Bayes algorithm with J48 decision tree algorithm for BC classification, and found that Naïve Bayes algorithm achieved higher accuracy. Hybrid model of Naïve Bayes algorithm and particle swarm optimization for BC classification has been proposed in [58], which achieved higher accuracy than Naïve Bayes algorithm alone. Authors of [59] applied Naïve Bayes algorithm with principal component analysis and linear discriminant analysis for breast cancer classification, and achieved an accuracy of 95.7%. An improved feature selection technique using Naïve Bayes algorithm for BC classification has been proposed in [60]. It achieved an accuracy of 96.4%.

Overall, Naïve Bayes algorithm has been shown to be an effective and efficient method for breast cancer classification, especially when combined with feature selection techniques and other optimization methods.

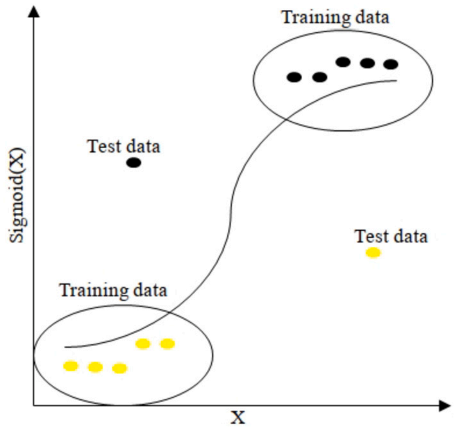

Fig. 5. Visual representation for LR method.

3.3. Logistic regression (LR)

Logistic Regression is a supervised binary classification algorithm. There could be only two outcomes for each sample. For BC investigation it can decide whether the tumor is benign or malignant [61,63]. Basically, the answer is always True (1) or False (0). Logistic regression is used in BC classification based on its several properties. It is effortless to upgrade the model with new datasets, it is less time consuming compared to Artificial Neural Networks (ANNs) and does not involve complex structures and lastly it is exceedingly adaptable for different categorization concepts [17]. Logistic regression uses an S shaped curve called sigmoid to fit the observed data and decide the class of the incoming data. In other words, LR is a sigmoid curve that ranges from 0 to 1, with the midpoint indicating the decision boundary between the two classes. Fig. 5 demonstrates the visual representation of the logistic regression.

Several studies have applied LR for BC classification. For example, a LR model with principal component analysis for BC classification has been proposed in [61], and achieved an accuracy of 94.6%. LR has been applied with a feature selection method in [62] for BC classification, and achieved an accuracy of 96.5%. A hybrid model of LR and particle swarm optimization for BC classification has been proposed in [63]. It achieved an accuracy of 97.2%.

Overall, LR has been shown to be an effective method for BC classification, especially when combined with feature selection techniques and optimization methods. However, LR assumes a linear relationship between the input features and the output variable, and may not capture complex nonlinear relationships that exist in the data.

3.4. Decision tree (DT)

Decision Tree is a supervised learning technique based on trees that can be applied to classification and regression problems. Classification trees are tree-like structures formed up of leaves that symbolize class tags and branches that represent distinct possibilities for those classes. To figure out the class of sample X, algorithm starts from the root node and tests the branches in the tree until it reaches the leaf node which contains the class prediction for the sample [18,64]. Fig. 6 displays a decision tree model as an illustration.

BC classification based on DT involves building a tree-like model of decisions and their possible consequences. Each node in the tree represents a feature, and the branches represent possible values of the feature. The leaves of the tree represent the class labels (i.e., malignant or benign). DT algorithm recursively splits the data into subsets based on the values of the features until the subsets are homogeneous with respect to the class labels.

Many recent studies have applied DT algorithms for BC classification. For example, authors of [66] proposed a decision tree-based approach for breast cancer classification that achieved an accuracy of 98.8%. A decision tree algorithm with feature selection for BC classification has been proposed in [67]. It achieved an accuracy of 98.5%. Another solution based on DT with feature selection and optimization for BC classification has been proposed in [68] and it achieved an accuracy of 97.3%.

DT algorithms have shown promising results for BC classification, especially when combined with feature selection and optimization techniques. However, decision trees can easily overfit to the training data and may not generalize well to new data. Therefore, more advanced machine learning algorithms such as random forests and gradient boosting may be more suitable for BC classification in certain cases.

3.5. Support vector machine (SVM)

Like DT, SVM is also a supervised learning algorithm which is useful for both regression and classification. It is a robust ML approach which uses both numerical and conceptual methods for the solution of regression problems and it is exceedingly precise for big datasets [19,65]. The main approach in SVM algorithm is to define a hyperplane (or decision surface) in a multidimensional space, a line if the space has only two dimensions, which will separate the labeled data points based on their classes. Since there can be more than one hyperplane to categorize data points, the best practice is to choose the line with the maximum margin which means providing the maximum possible distance between two data sets to offer more reliable classification. Support vectors, as the name implies, are vital in determining the hyperplane; these are the data points that are closest to the decision surface. The concept of the SVM algorithm is visualized in Fig. 7 so that it can be understood more profoundly.

SVM is a popular algorithm for breast cancer classification due to its ability to handle high-dimensional data and nonlinear relationships between variables. Recent studies have utilized SVM for BC classification and achieved high accuracy. For example, Zou et al. [69] proposed a hybrid feature selection and SVM approach for breast cancer classification, achieving an accuracy of 98.9%. In [70], authors proposed a novel kernel function and SVM-based approach for BC classification, achieving an accuracy of 96.85%. In the study [71] a feature selection and SVM-based approach for BC classification has been proposed and it achieved an accuracy of 97.78%.

SVM is a powerful algorithm for BC classification that can achieve high accuracy when combined with appropriate feature selection, kernel function, and optimization techniques.

3.6. Random forest algorithm (RF)

RF is a supervised learning method which uses multiple decision trees as a resolving mechanism [72,73]. It deals with classification and regression problems. It reaches a solution by merging the outcomes of several decision trees. It is employed to eliminate the disadvantages of decision tress like over fitting. Considering other classification algorithms, RF offers consistency and certainty by decreasing total error rate of the classification. Furthermore, it is an efficient tool to manage unbalanced data [20]. Fig. 8 shows the graphical form of random forest algorithm.

RF is a popular algorithm for breast cancer classification because it can handle high-dimensional data, nonlinear relationships between variables, and is able to avoid overfitting. Recent studies have utilized Random Forest for breast cancer classification and achieved high accuracy. Many solutions have been proposed based on RF. For example, Riaz et al. [72] proposed a feature selection and RF-based approach for BC classification that achieved an accuracy of 98.05%. The study [73] proposed a hybrid feature selection and RF-based approach for BC classification that achieved an accuracy of 97.06%. A RF-based approach with an optimized feature selection method for BC classification has been proposed in [74] and it achieved an accuracy of 98.95%.

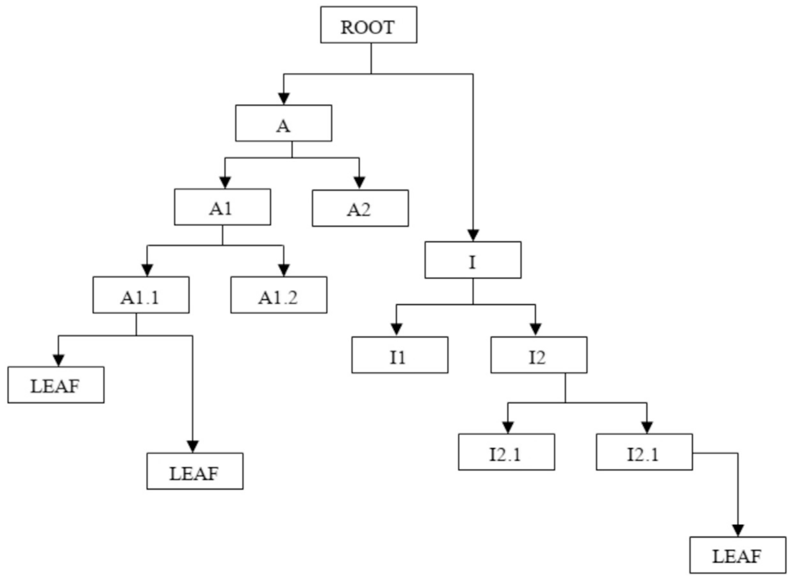

Fig. 6. A sample decision tree for a better understanding about the concept showing root and leaf nodes and branches.

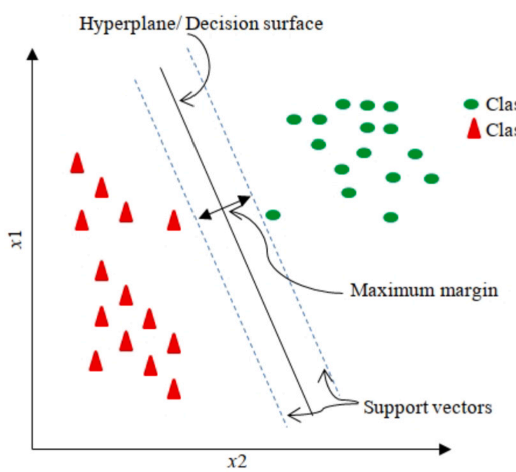

Fig. 7. Visualization of SVM algorithm.

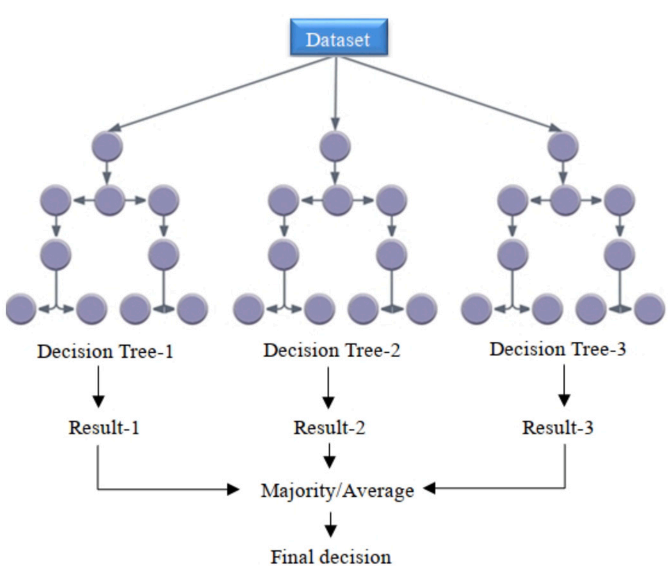

Fig. 8. Graphical demonstration of random forest method.

Basically, RF is a powerful algorithm for BC classification that can achieve high accuracy when combined with appropriate feature selection, parameter tuning, and optimization techniques.

4. ML for BC studies

Modern technologies such as ML, AI and DL have recently proven effectiveness in medical research [7,19,25,34,36,37,47]. With the evolution of screening systems and ML methods, digital histopathology image analysis provides cheap and reliable BC diagnosis options [21]. In addition, Image processing combined with ML is also being used to interpret mammograms, sonograms or MRI scans for the detection of abnormal formations inside the breast. Though, the majority of studies are focused on the diagnosis and prognosis of BC. There are also some studies that focus on survivability prediction. Kate and Nadig, for example, have used SEER data set to compare stage-specific and traditional prediction models by using logistic regression, decision Tress and naïve Bayes methods to estimate survivability of female patients [22]. This section will focus on research that use ML algorithms to classify BC based on screening images and histological analysis, as well as predictive models for survival and relapse rates.

4.1. *Screening*

X-Ray Mammography, MRI and Ultrasound are the most common screening systems used to detect BC. Despite the fact that these systems are less expensive, faster, and more effective at diagnosing cancer, understanding the resulting images are highly depended on humans. As a result, the accuracy of the diagnosis can be easily influenced by the physical or psychological condition as well as the experience level of the radiologist. Thus, Computer Aided Diagnosis (CAD) systems have started to gain popularity as a tool to assist doctors. However, they still require improvements. This improvement can be made based on ML. Fig. 9 illustrates a sample BC for each imaging methods discussed.

4.1.1. *Mammography*

Although it has some draw backs like false positive and radiation exposure, mammography has been shown to reduce BC patient fatalities by almost 20% and is still considered as the golden standard for the diagnosis of BC due its high susceptibility, cost-efficiency and precision [23]. It is commonly used in routine controls for women who are over 40 because it does not function well in dense breast tissues which are generally found in younger women. Mammography is also useful for

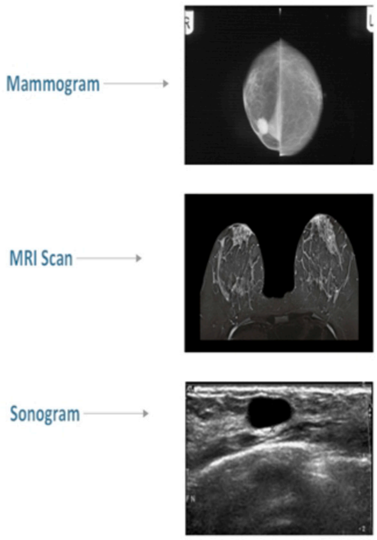

Fig. 9. Breast scans gathered with mammography, MRI and ultrasound.

tracking the progress of suspicious lumps. The first step towards cancer classification through mammograms is pre-processing which includes image enhancement. Enhancing involves increasing the contrast of the images to provide a more detailed mammogram scan [24]. The process continues with segmentation, feature extraction, feature selection and classification. Segmentation means dividing the image to several parts in order to get more accurate results, then certain characteristics are extracted and picked with the aim of categorizing the breast scan as benign or malignant [25]. Fig. 10 demonstrates the required steps to implement a mammogram classification algorithm.

There are lots of popular and well-known algorithms that are being utilized for the mentioned systems like KNN, SVM or CNN which are supervised methods. On the other hand, there are also studies which compare supervised (original) and semi-supervised GrowCut algorithm to classify tumors in mammograms. Even though the studies clearly demonstrate that the supervised GrowCut algorithms resulted in more accurate classification, the semi-supervised models provided advantages such as being independent of a specialist's supervision [26]. And for the conventional algorithms, Kaur et al. [27] analyzed and compared the results of Multi-Class SVM (MSVM) algorithm combined with K-mean Clustering to DT algorithm and realized that MSVM offers more precision and consistency.

4.1.2. MRI

Unlike mammography, MRI is a more time-consuming application which results in higher specificity and fewer false-negative or false-positive. As a consequence, it is generally recommended for women who are at high risk of BC. The accuracy of the results is independent of the patient's age or breast density, making it more consistent. However, similar to mammograms MRI results are hard to analyze too. Unfortunately, the lack of publicly accessible free data sets has resulted in a scarcity of MRI classification studies [28]. While reading an MRI scan, radiologists use a 7 stages assessment procedure called BI-RADS. BI-RADS 0 suggests that the data is not sufficient to make a statement and additional evaluations are needed BI-RADS 1 means the scan is negative for any malignant formation, and BI-RADS 6 means definite malignancy. This assessment is based on characteristics of the suspicious tissue. These futures can be listed as: shape, margins and internal enhancement characteristic of masses. BI-RADS procedure is subjective and heavily reliant on the radiologists' experience and success. Consequently, Zhou et al. [29], utilized 3D deep CNN to diagnose malignant tumors and lesions seen in dynamic contrast enhanced (DCE) MRI scans with an accuracy of 83.7 percent and a sensitivity of 90.8 percent. The data set used in the research included scans of 1537 women and the scans were labeled by radiologists based on the BI-RADS standards [29].

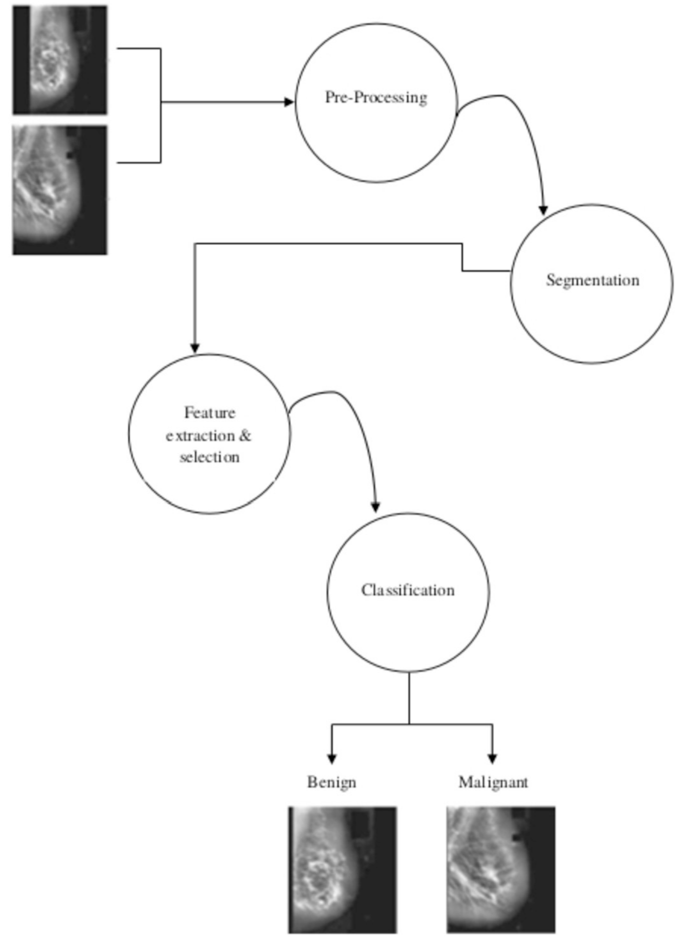

Fig. 10. These are the necessary steps to take while constructing a mammography classifier system.

4.1.3. Ultrasound

Ultrasound creates images of the breast tissue that are called sonograms by sending and receiving sound waves. In spite of the fact that a lot of researchers oppose the usage of ultrasound in BC diagnosis because the routine procedure suggests mammograms, ultrasound is good to detect any palpable or impalpable masses inside the breast tissue. Moreover, it does not emit ionized radiation, ensuring that the patient is not exposed to harmful screening. The algorithms for ultrasound image classification mainly include 4 parts. Pre-processing to enhance the contrast and reduce noise, segmentation to determine regions of interest (ROIs), future extraction and selection which basically involves defining the specific indicators for benign/malignant tissue, and finally classification which can be described as categorizing the suspicious parts of the breast as benign or malignant [30]. DT, Bayesian Networks, SVM and Linear classifiers are the most commonly used algorithms in this type of models.

4.2. Pathology

Histopathological images are created by the procedure of analyzing the biopsy sample taken from the patient with a light microscope after using hematoxylin for blue color and eosin for contrasting shades to stain the nuclei and cytoplasmic non-nuclei structures in the cell respectively [31]. For every type of cancer histopathological look differ from each other. Fig. 11 exhibits histopathologic slide samples for previously introduced cancer types. Pre-processing is the first stage in assessing

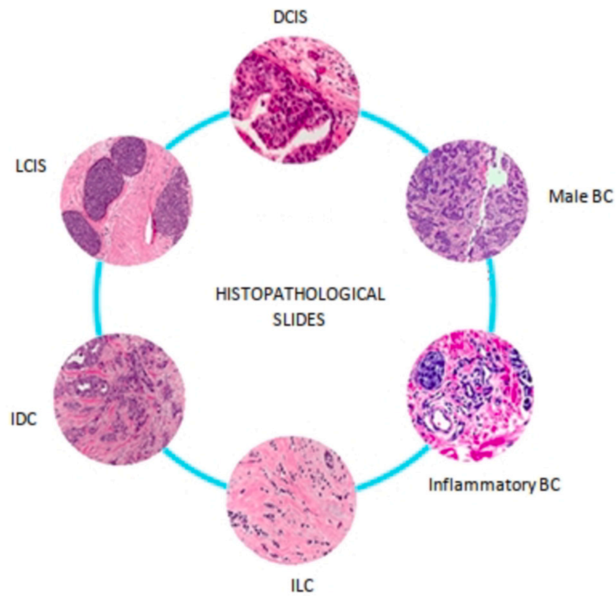

Fig. 11. Samples of histopathological analysis for different types of breast cancer: DCIS, LCIS, IDC, ILC, Inflammatory BC and Male BC.

histopathologic data. It is a significant stage because pathology analysis of breast tissue occasionally has low resolution and high noise which may result in deficiencies in the outcomes of the classification algorithm. Pre-processing of histopathological slides that can include balancing the data amount in each sample class to prevent wrong sorting decisions, enhancing the contrast of the image as in mammography studies, re-segmenting the images by turning pixels to super-pixels or normalizing the colors in images [31]. There are two viable classification methods for BC classification of histopathological images: binary and multiclass classification. Binary classification has two possible outcomes which are true or false implying that the model can only answer whether the biopsy sample is malignant or not, whereas multiclass classification allows for learning the sub-type of the cells [32,33].

4.3. BC classification based on DNN and transfer learning

SVM and probabilistic neural networks to identify cell nuclei and classify BC have been used in [35]. Breast cytological pictures were used for the experiments, and the resulting error rate, correct detection rate, sensitivity, and specificity were used to compare the outcomes. According to their individual technique, the results obtained are more effective and may be used to multiple datasets. In-depth research on the categorization of breast cancer using DL and traditional ML techniques was given in [36]. To distinguish between benign and malignant lesions, they extracted features from the images using the color histogram and Haralick textures. Their suggested strategy produced accuracy results that ranged from 93.25% to 93.97% [36].

A thorough review of the utilization of ML and DL applications in breast cancer diagnosis was given in [37]. They went into great detail on the studies and literature review pertaining to the categorization of breast cancer. Additionally, they emphasized the advantages and disadvantages of these strategies. By incorporating the findings of earlier research, the authors of this study came to the conclusion that deep learning approaches are far more appropriate for the task of classifying images of BC when the datasets are larger [37]. A thorough and in-depth evaluation of deep learning and ML methods for BC classification and diagnosis using medical imaging was provided in [53]. They demonstrated the plethora of innovative applications utilized in medical diagnosis as well as the quick adoption of deep and ML in the area [53].

In [38] authors suggested applying ML-based models to the challenge of classifying BC. They asserted that the accuracy of breast cancer detection and classification achieved by physicians on average is approximately 79%, but the accuracy of their suggested model is 91% [38]. Authors of [39] classified BC pictures using 569 samples and 30 characteristics using the Wisconsin Breast Cancer Dataset. The Kaggle repository is where the dataset was gathered. They used precision and accuracy as metrics to gauge how effective their job was. They used the artificial neural network K Nearest Neighbor, logistic regression, and SVM as their methods. To achieve the desired outcomes, they have each used them independently. They consequently accomplished a maximum average accuracy of 99.3% for the task of classifying images of BC [39,40].

A multi-deep CNN was proposed in the study [41]. for the categorization of breast cancer. They used DNN that had already been trained to extract information from the images. The same features were also employed by the SVM classifier with various kernel functions. To further minimize the feature vector, the research study's authors employed principal component analysis. When compared to other cutting-edge CAD systems, they asserted that their results were the best. Using deep learning, authors of [42] demonstrated an effective method for classifying skin cancer. They used the real-time dataset that was gathered from the Pakistani DHQ in Faisalabad for this purpose. They divided the images of skin cancer into categories, such as melanoma and non-melanoma. Their findings showed a 93.29% classification accuracy.

In [43] authors classified the BC images into benign and malignant tumors using the concepts of transfer learning and deep learning. Using pre-trained GoogLeNet, VG-GNet, and ResNet architectures, they extracted the features. Additionally, the average pooling idea was incorporated to these features in the fully connected layers. They employed a total of 8000 images as their dataset; of these, 6000 images were used for training the network, while the remaining 2000 images were used for testing. Their suggested method's highest average accuracy was reported to be 97.25%. In their research, authors of [44] suggested classifying the human charred skin images using a CNN. There were four hundred and fifty images in their dataset. 35% of the images were used to assess the accuracy of the network after it had been trained with 65% of the total. They classified the images with 79.4% accuracy as a result. They expanded the dataset, which included 600 images, in a different study. The segmentation and classification of burnt human skin was done using the same methodology. The classifier's accuracy was trained and tested using a 60% and 40% ratio. They classified the images with 83.4% accuracy as a consequence [45].

Images of BC can be classified using ensembles of deep learning models, as demonstrated in [46]. Four distinct models were trained using the pre-trained VGG16 and VGG19 architectures. Their total classification accuracy of 95.29% for the various kinds of BC. Various supervised ML methods to classify the BC images into malignant and non-malignant tumors have been used in [47]. K-Nearest Neighbor, Logistic Regression, Decision Tree, Random Forest, and SVM were their methods. Additionally, they employed Adam Gradient Descent Learning models, with a 98.24% accuracy rate.

A thorough investigation of the use of deep learning and imaging techniques for the classification of BC images was provided in [28]. They selected 49 distinct research papers from 8 separate archives that had been published in journals and conferences. They outlined every ML and DL method that is applied to various medical picture classification tasks, such as the classification of BC. For novice researchers, they showcased the 10 open research challenges, one of which being the automatic identification of BC. Bidirectional Encoder Representations (BER) to show the clinical data of BC patients has been used in [48]. Additionally, they extracted characteristics or attributes from the infected images of breast cancer patients using deep learning algorithms. The greatest outcomes attained with the aforementioned strategies were noted as 96.73%.

In their research on the identification and categorization of BC, authors of [49] introduced a mathematically realized DL assisted Efficient Adaboost Algorithm (DLAEABA) with the capacity for advanced level of calculations. Additionally, they employed a few computer vision techniques. Additionally, they provided a comparison by displaying the 97.2% accuracy of the experimental results. To detect cell nuclei, [50] presented an understandable method with certain fuzzy logic approaches features. For the purpose of classifying and segmenting NC, deep CNN was employed. Their suggested approach required less computing time and was far more accurate. To verify the accuracy of the classification process, they employed the tenfold cross-validation approach. In their suggested research project, they demonstrated an accuracy of 98.62%.

The literature review conducted in this section provides a detailed analysis, which indicates a significant deficiency in the early diagnosis of cancer utilizing deep and transfer learning methods, resulting in low accuracy. The research indicates that undertaking the activity can be challenging due to limited resources or a shortage of skilled and knowledgeable professionals. Considerable research has been conducted by scholars in the medical domain; yet, there are deficiencies in the precision of their findings. In order to address these issues, efforts have been made to enhance the accuracy and effectiveness of breast cancer picture classification. This has been achieved through the integration of deep learning and transfer learning methodologies, aiming to enable accurate and promising early-stage diagnosis of breast cancer.

Table 1 gives a summary about some classification methods of BC.

5. IoT and 5G technologies in BC

IoT (Internet of Things) can be utilized in various ways to enhance BC detection, treatment, and patient care [75–78]. Here are some ways IoT can be applied in the context of breast cancer:

- Early Detection and Screening:
  - Smart Bras and Wearables: IoT-enabled bras or wearables with embedded sensors can monitor changes in breast temperature, moisture, or tissue density. These changes could potentially indicate early signs of breast abnormalities, prompting women to seek further medical evaluation.
  - Remote Mammogram Monitoring: IoT devices can transmit real-time data from mammography machines to radiologists, allowing for remote monitoring and quicker analysis of mammogram results.
- Treatment and Medication Management:
  - Smart Drug Dispensers: IoT-connected medication dispensers can help patients manage their breast cancer treatment regimen by providing reminders, tracking adherence, and notifying healthcare providers of missed doses.
  - Wearable Vital Sign Monitors: Wearable devices can continuously monitor vital signs such as heart rate, blood pressure, and temperature. Any significant deviations from normal values could trigger alerts to healthcare providers or patients, indicating potential complications or side effects of breast cancer treatments.
- Post-Surgery Recovery and Rehabilitation:
  - IoT-Enhanced Prosthetics: IoT sensors integrated into prosthetic breasts can help patients recover and adapt to post-mastectomy changes more comfortably. These sensors can monitor skin health, detect pressure points, and provide feedback on fit and comfort.
  - Remote Rehabilitation: IoT devices can assist with remote physical therapy and rehabilitation programs, ensuring that breast cancer survivors receive appropriate care and guidance to regain strength and mobility.
- Patient Support and Education:
  - Smart Apps and Chatbots: Mobile apps and chatbots can provide patients with personalized information about breast cancer, treatment options, and post-treatment care. They can also offer emotional support and connect patients to support groups.
  - IoT-Enabled Telemedicine: Telemedicine platforms integrated with IoT devices allow patients to consult with healthcare professionals remotely, reducing the need for frequent in-person visits during treatment.
- Data Collection and Research:
  - Wearable Research Devices: IoT-enabled wearables can collect anonymized health data from breast cancer patients, contributing to large-scale research studies and clinical trials. This data can help researchers identify patterns, predict outcomes, and develop more effective treatments.
- Monitoring Disease Progression:
  - IoT Imaging: Advanced imaging equipment with IoT capabilities can monitor changes in breast tissue over time, aiding in the assessment of disease progression or treatment effectiveness.
- Preventive Measures:
  - Environmental Sensors: IoT sensors can monitor environmental factors that may contribute to breast cancer risk, such as pollution or exposure to harmful chemicals. This data can inform public health initiatives and help individuals make informed decisions about their surroundings.

It is important to note that while IoT offers significant potential benefits in breast cancer care, privacy and security concerns must be addressed to protect sensitive patient data. Additionally, these IoT applications should always be used in conjunction with traditional medical assessments and under the guidance of healthcare professionals.

5G technology has the potential to significantly enhance the detection and treatment of breast cancer in various ways due to its high-speed, low-latency, and high-capacity capabilities [79–84]. Here are some ways in which 5G technologies can be leveraged for breast cancer detection and management:

- Real-Time Medical Imaging:
  - High-Resolution Imaging: 5G can enable real-time transmission of high-resolution medical images, such as mammograms, MRIs, and ultrasounds, to radiologists and specialists, allowing for quicker and more accurate diagnoses.
  - Remote Consultations: Surgeons and radiologists can use 5G-powered telemedicine platforms for remote consultations and collaborations, providing expert guidance even in remote or underserved areas.
- IoT-Enhanced Medical Devices:
  - IoT Sensors: Medical devices equipped with IoT sensors can transmit data in real-time over 5G networks. For instance, smart biopsy needles or ultrasound probes can provide precise and immediate feedback during procedures.
  - Remote Monitoring: Patients undergoing breast cancer treatment can wear IoT-connected devices that continuously monitor vital signs and treatment-related parameters. This data can be transmitted to healthcare providers via 5G for real-time monitoring and intervention.
- AI and Machine Learning:
  - Faster AI Processing: 5G can accelerate the processing of AI algorithms used for breast cancer detection and analysis. AI can assist radiologists by quickly identifying anomalies in medical images, leading to faster and more accurate diagnoses.
  - Predictive Analytics: With 5G, AI models can analyze large datasets of patient information and imaging data to predict breast cancer risk, disease progression, and treatment outcomes more effectively.
- Augmented Reality (AR) and Virtual Reality (VR):

Table 1
Summary of relevant researches BC classification.

Ref	Year	Method	Dataset	Feature selection	Feature extraction	Result and accuracy
[56]	2013	KNN	WBCD	-	-	In this study, two types of distance that gave the best results 98, 70% and 98, 48% for Euclidean distance and Manhattan respectively.
[59]	2015	NB, weighted-NB	WBCD	-	-	In this study the weighted-NB can reach an accuracy of 98.54%.
[85]	2017	SVM, KNN	MIAS	-	GLCM	SVM accuracy is 95.7%. It is higher than KNN.
[86]	2017	BF, SVM, SLR, NB, KNN, AdaBost, Fuzzy, DT-J48	699 instances-UCI	Probability technique	-	SVM gave the highest accuracy (97.07%).
[87]	2018	KNN, NB	683 samples-UCI	Probability technique	-	KNN has an accuracy of 97.51% in this study.
[58]	2018	NB, KNN, DT	WBCD	Based on PSO (Particle Swarm Optimization)	-	Without PSO feature selection, DT classifier provided the highest accuracy level of 76.3%. However, when PSO feature selection was used, the best result of the accuracy level was exhibited by NB (81.3%).
[88]	2018	ANN, SVM, KNN	250 mammogram DDSM	T-test algorithm	Wavelet transform	In this study ANN achieved an accuracy of 98.9%. It was higher than SVM and KNN.
[89]	2018	KNN	120 images Mini-MIAS	-	GLCM	92% of accuracy has been saved by KNN on this dataset.
[90]	2018	SVM, DT, NB, KNN	357 dataset WBCD	features that are calculated from a digitized image of a fine needle aspiration (FNA)biopsy	-	SVM got an accuracy of 92% on the selected dataset of this study. It was gather than DT, NB and KNN.
[91]	2019	SVM, KNN	30 images-DDSM	-	Contrast Correlation	In this study, KNN has the highest accuracy, which was 97%.
[92]	2019	KNN, ANN	50 CESM images BAHAYA foundation	-	Curve fitting	ANN has an accuracy of 92% and KNN achieved a higher accuracy (96%).
Ref	Year	Method	Dataset	Feature selection	Feature extraction	Result and accuracy
[93]	2020	KNN, LR, EL	569 instances-WDBC	Statical technique correlation coefficient	PCA	EL (ensemble learning) has the highest accuracy in this study. It was 99.30%
[94]	2020	KNN, DT, SVM, Adaboost	569 instances-WDBC	NCA	-	99.12% has been achieved by KNN.
[95]	2020	SVM, DT, NB, LR, LDA, KNN	569 instances - WDBC	-	PCA	In this study, LR reached an accuracy of 97.23% and LDA had an accuracy of 95.73%. However, SVM had the highest accuracy of 98%.
[36]	2020	VGG16, VGG19, ResNet50	BreakHis	-	Pre-trained Network	Based on this study, the combination VGG16 + SVM provides the best result (80% of accuracy).
[60]	2021	SVM, KNN, NB, perceptron ML, LR	WBCD	Univariate selection	Recursive feature elimination	In this study, the accuracy of SVM, KNN, Naive Bayes, Perceptron and LR with ten features are 98.8%, 98.357%, 96.985%, 98.049%, 98.968% respectively. However, with feature reduction, LR has the best accuracy (99.968%).
[96]	2022	DRDA-Net	BreakHis	-	-	The proposed model showed acceptable accuracy (98.1%). Densely connected blocks addressed the overfitting and vanishing gradient problems.
[97]	2022	CNN	BreakHis	-	DenseNet201 and VGG16 architecture models	In this study, the proposed model reached 100% accuracy.
[98]	2022	CNN	Wide-scale data	-	-	The proposed DL architecture could achieve a good average of accuracy 98%.
[74]	2023	SVM, KNN, NB, EL, DT	WBCD, MBCD	LASSO	-	SVM showed the highest accuracy for WBCD and MBCD (97.95%, 98.95%).

  - Surgical Assistance: Surgeons can use AR and VR technologies powered by 5G to enhance the precision of breast cancer surgeries. These technologies can provide real-time guidance, overlaying critical information on the surgeon's field of view.
- Data Sharing and Collaboration:
  - Data Integration: 5G facilitates the secure and rapid sharing of patient data among healthcare professionals and institutions. This can lead to better-informed treatment decisions and coordination of care.
- Remote Pathology and Tumor Board Meetings:
  - Pathology Collaboration: Pathologists can use 5G to share digital pathology slides in real-time, enabling remote consultations and multidisciplinary tumor board meetings for more accurate diagnoses and treatment planning.
- Clinical Trials and Research:
  - Remote Monitoring for Clinical Trials: 5G allows for remote monitoring of patients participating in breast cancer clinical trials. Researchers can collect data in real-time, enhancing the efficiency of trials and reducing the need for physical visits.
- Patient Education and Support:
  - High-Definition Patient Education: 5G can support high-quality video content for patient education and support, helping individuals understand their diagnosis, treatment options, and post-treatment care.

While 5G technologies hold great promise for breast cancer detection and management, it's essential to address privacy and security concerns related to sensitive patient data transmitted over these networks. Additionally, regulatory and ethical considerations must be taken into account when implementing these technologies in healthcare settings.

6. Future trends and challenges

ML studies for BC have evolved in recent years, as described in the previous sections. A vast number of studies have been published in the literature based on ML algorithms with different imaging modalities such as MRI, ultrasound and mammography to classify BC type. There are also numerous studies on categorizing the histopathological data to determine the molecular sub-types of the tumors and using ML models to predict the recurrence in cancer patients and 5-years survival rate. The capabilities of the used models and algorithms to assist medical practitioners in classification have resulted in an increase in the use of AI approaches for BC [34]. However, there are significant challenges that prevent these models from being fully implemented as a gold standard in clinical practice. This section will go through some of the challenges and the future steps that needed to be taken to overcome those.

6.1. Enhanced datasets

Lack of convenient datasets in BC classification for both medical images and histopathological reports is a major challenge for ML studies. There are couples of aspects that need to be satisfied by datasets such as volume, diversity, availability and trustworthiness to provide proper usage. One of the main drawbacks of implementing ML techniques to medical image classification is that these methods are vastly depended on supervised learning algorithms [51]. Due to the fact that supervised learning methods in radiology studies require large sizes of image datasets, which are formerly labeled by radiology experts to accomplish better accuracy [51]. Unfortunately, currently used BC datasets are relatively small. Moreover, as it was stated by the anthers of [52].

Datasets available to public use including MRI, Mammography and Ultrasound images are less in number and unbalanced in terms of providing similar amounts of cancerous-benign objects and divergence in pathology. A reason for the lack of mass datasets is that, despite the large amount of medical data generated in hospitals and clinics everyday identifying the useful material is a delicate process that necessitates a great deal of effort. Besides, another struggle is to maintain absolute accuracy in annotation without any false labeled data. Annotation is a labor-intensive and time-consuming process because it requires a second opinion from a couple of medical specialists in order to eliminate human errors [52], [53]. It is also challenging to adapt classification with ML to currently emerging imaging modalities such as breast thermography or microwave imaging. Since these methods are not standard clinical practice the amount of available data is highly restricted [54].

To get over the constraints caused by insufficient datasets, official attempts by governments or healthcare organizations must be made the collect of the unused data from various medical centers like personal clinics or hospitals. The collected data must be processed and annotated with a large workforce, to be utilized in ML models. Even though it seems to be too much hassle at the moment, with ML algorithms showing to be efficient in BC diagnosis, studies in building cancer datasets are remarkably potential to accelerate.

6.2. Improved algorithms

KNN, Naïve Bayes, LR, DT, SVM, RF and DNN are the most commonly used algorithms. The goal of employing these algorithms or other ML models in cancer classification is to eliminate human error and achieve 100% accuracy. Though the majority of these methods show promising outcomes, the accuracy of the algorithms still needs enhancements in order for these techniques to replace the medical specialists in BC diagnosis and prognosis [51]. Putting accuracy, a side, the specificity and consistency are also crucial features that require enhancements. The efficiency of the algorithms must not change according to different datasets or various types of cancer. If ML models are to be standardized for routine clinical practice, the algorithms must perform with a high level of accuracy and precision in every BC patient without the need of human intervention. With the huge developments in AI and DL studies, improvements in ML algorithms for BC studies are predicted to occur shortly.

6.3. Defining standards and clinical practice

While ML studies for healthcare industry are rapidly progressing, another concern is how to define valid standards so that these cutting-edge technologies can be integrated into ordinary clinical practice. The authors of [34], suggested to generate and use standardized public datasets with medical images from several systems in the forthcoming applications. Other than the standardized datasets, annotation methods and the use cases which are coherent with legal obligations and ethical concerns must also be standardized by responsible organizations [55].

After the determination of the regulations, the next step is employing ML models to clinical practice. This stage requires interdisciplinary work between pathologists, radiologists, cancer specialists and computer scientists. To begin, the models must be tested in several pilot areas with different type of patients under the supervision of medical experts so that the accuracy of the models can be enhanced with feedbacks. Later on, assuming that the models performed as expected in terms of accuracy and specificity in the trial areas, they can be used in general healthcare.

6.4. Potential research directions

This section will outline some factors, which can be seen as potential research directions in the field of BC that are based on AI and ML. When considering methods to improve the diagnosis of BC, it is important to take into account the following factors:

- Creating new datasets is necessary due to the limited availability of extensive and trustworthy datasets from experienced physicians for the diagnosis of BC.

- There is a limited supply of pre-trained models specifically designed for medical images, particularly BC. This necessitates the need to evaluate and test novel models that utilize AI and ML. In addition, these models must be available.
- Employing several imaging modalities and numerous modalities for the categorization of BC, which encompasses DBT, US, and MRI images.
- Further investigation is required to create CAD models that make use of unlabeled images, which are a crucial source of information.
- It is crucial to integrate all the knowledge acquired from different stages while considering its validity period. Additional study is required to accurately forecast the survival rates of BC patients based on the stage and year of diagnosis.
- Integrating non-imaging information with imaging data is a beneficial approach. Integrating radiomic characteristics with image data can improve the performance of AI and ML models.
- AI and ML models must also include the class imbalance problem, which refers to the unequal number of positive and negative data.

7. Discussion

The utilization of AI technology in imaging diagnosis has great potential in the field of BC detection. There are numerous appealing possibilities for applying AI in various areas. For instance, AI-driven diagnosis refers to tests that are minimally invasive or non-invasive and provide convenient and effective assessment of efficacy. In addition, AI helps differentiate the unique molecular biological characteristics of primary tumors and recurrent/metastatic sites. AI enables the acquisition of information about tumor heterogeneity. Furthermore, AI aids in identifying treatment response or tumor progression, particularly in the context of immunotherapy. Nevertheless, the existing methods, ranging from imaging detection techniques to lesion segmentation and qualitative assessment of images, are insufficient and inadequate to facilitate the utilization of AI-assisted imaging diagnosis for independent imaging or clinical diagnosis. There are several significant factors that restrict the widespread use of AI-assisted imaging diagnosis in BC. Firstly, there is an absence of a universally acknowledged operational standard for the procedure, encompassing imaging detection, tumor segmentation and image feature extraction, in order to guarantee consistent and replicable outcomes. Additional, innovative algorithms with more specificity are required to handle many types of images with different levels of quality, as well as variances specific to individual patients. Furthermore, while imaging-based AI diagnosis holds promise for reducing the elevated false positive rate in BC detection, its diagnostic accuracy is contingent upon the quality of imaging detections such as mammography and ultrasound. Additionally, the algorithm used in this process is still in need of refinement to enhance its diagnostic accuracy. Moreover, the clinical utility of AI-assisted BC diagnosis must be confirmed through a randomized clinical trial involving a large and diverse group of patients.

8. Conclusion

This survey underscores the critical role of AI, particularly ML DL techniques, in advancing the diagnosis of BC. The exploration of recent CAD systems reveals a notable shift towards more sophisticated models that leverage the power of AI for improved accuracy and efficiency in medical image analysis. The comparison with established methods not only highlights the progress made but also underscores the technical nuances and trade-offs associated with each model. Additionally, the integration of IoT devices in BC research and treatment emerges as a transformative force, enabling real-time monitoring and personalized healthcare solutions. Wearable sensors and smart implants, coupled with the capabilities of 5G technology, promise to revolutionize communication speeds and data transfer in the realm of BC applications. This connectivity enhancement facilitates remote diagnostics, telemedicine, and collaborative research efforts, ultimately contributing to the acceleration of innovation and improvement in patient outcomes.

The study emphasizes the superiority of DL over traditional ML methods, especially in scenarios involving extensive datasets, highlighting the need for ongoing research to bridge existing knowledge gaps. The identified gaps underscore the urgency for both practical and scientific investigations to further enhance healthcare outcomes. In conclusion, the combined integration of advanced CAD models, IoT devices and 5G technology holds immense promise for the future of BC detection, treatment, and patient care, urging continued exploration and investment in these transformative technologies.

Declaration of competing interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Acknowledgements

This research was supported by Istanbul Technical University (ITU). Meanwhile, this work was supported by the Higher Institution Centre of Excellence (HICOE) program, Ministry of Higher Education (MOHE) Malaysia, conducted at the Universiti Teknologi Malaysia under the HICOE Research Grant R.J130000.7823.4J637. Also, this study is supported via funding from prince Sattam bin Abdulaziz University project number (PSAU/2023/R1444).

References

[1] WHO, https://www.who.int/news-room/fact-sheets/detail/breast-cancer. (Accessed 4 January 2023), 2023.
[2] https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/estimated-number-of-new-cancer-cases-and-deaths-by-sex-2020.pdf. (Accessed 4 January 2023), 2023.
[3] WCRF, https://www.who.int/news-room/fact-sheets/detail/breast-cancer. (Accessed 4 January 2023), 2023.
[4] R. Wang, Y. Zhu, X. Liu, X. Liao, J. He, L. Niu, The clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer, BMC Cancer 19 (1) (2019) 1–12.
[5] Machine Learning, https://www.ibm.com/cloud/learn/machine-learning. (Accessed 4 January 2023), 2023.
[6] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Comput. Struct. Biotechnol. J. 13 (2015) 8–17.
[7] Y. Li, Z. Chen, Performance evaluation of machine learning methods for breast cancer prediction, Appl. Comput. Math. 7 (4) (2018) 212–216.
[8] F.F. Ting, Y.J. Tan, K.S. Sim, Convolutional neural network improvement for breast cancer classification, Expert Syst. Appl. 120 (2019) 103–115.
[9] M.A. Mohammed, B. Al-Khateeb, A.N. Rashid, D.A. Ibrahim, M.K. Abd Ghani, S.A. Mostafa, Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images, Comput. Electr. Eng. 70 (2018) 871–882.
[10] R. Roslidar, A. Rahman, R. Muharar, M.R. Syahputra, F. Arnia, M. Syukri, et al., A review on recent progress in thermal imaging and deep learning approaches for breast cancer detection, IEEE Access 8 (2020) 116176–116194.
[11] Y. Yari, T.V. Nguyen, H.T. Nguyen, Deep learning applied for histological diagnosis of breast cancer, IEEE Access 8 (2020) 162432–162448.
[12] J.A. Cruz, D.S. Wishart, Applications of machine learning in cancer prediction and prognosis, Cancer Inform. 2 (2006) 117693510600200030.
[13] M. Akram, M. Iqbal, M. Daniyal, A.U. Khan, Awareness and current knowledge of breast cancer, Biol. Res. 50 (2017) 1–23.
[14] A. Prat, E. Pineda, B. Adamo, P. Galván, A. Fernández, L. Gaba, et al., Clinical implications of the intrinsic molecular subtypes of breast cancer, Breast 24 (2015) S26–S35.
[15] K.L. Maughan, M.A. Lutterbie, P.S. Ham, Treatment of breast cancer, Am. Fam. Phys. 81 (11) (2010) 1339–1346.
[16] Y. Ruan, X. Xue, H. Liu, J. Tan, X. Li, Quantum algorithm for k-nearest neighbors classification based on the metric of Hamming distance, Int. J. Theor. Phys. 56 (2017) 3496–3507.

[17] L. Khairunnahar, M.A. Hasib, R.H.B. Rezanur, M.R. Islam, M.K. Hosain, Classification of malignant and benign tissue with logistic regression, Inform. Med. Unlock. 16 (2019) 100189.
[18] H. Sharma, S. Kumar, A survey on decision tree algorithms of classification in data mining, Int. J. Sci. Res. 5 (4) (2016) 2094–2097.
[19] N. Fatima, L. Liu, S. Hong, H. Ahmed, Prediction of breast cancer, comparative review of machine learning techniques, and their analysis, IEEE Access 8 (2020) 150360–150376.
[20] M. Zhu, J. Xia, X. Jin, M. Yan, G. Cai, J. Yan, G. Ning, Class weights random forest algorithm for processing class imbalanced medical data, IEEE Access 6 (2018) 4641–4652.
[21] X. Li, M. Radulovic, K. Kanjer, K.N. Plataniotis, Discriminative pattern mining for breast cancer histopathology image classification via fully convolutional autoencoder, IEEE Access 7 (2019) 36433–36445.
[22] R.J. Kate, R. Nadig, Stage-specific predictive models for breast cancer survivability, Int. J. Med. Inform. 97 (2017) 304–311.
[23] S.H. Jafari, Z. Saadatpour, A. Salmaninejad, F. Momeni, M. Mokhtari, J.S. Nahand, et al., Breast cancer diagnosis: imaging techniques and biochemical markers, J. Cell. Physiol. 233 (7) (2018) 5200–5213.
[24] S. Anand, S. Gayathri, Mammogram image enhancement by two-stage adaptive histogram equalization, Optik 126 (21) (2015) 3150–3152.
[25] G. Meenalochini, S. Ramkumar, Survey of machine learning algorithms for breast cancer detection using mammogram images, Mater. Today Proc. 37 (2021) 2738–2743.
[26] F.R. Cordeiro, W.P.D. Santos, A.G. Silva-Filho, Analysis of supervised and semi-supervised GrowCut applied to segmentation of masses in mammography images, Comput. Methods Biomech. Biomed. Eng. Imaging Vis. 5 (4) (2017) 297–315.
[27] P. Kaur, G. Singh, P. Kaur, Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification, Inform. Med. Unlock. 16 (2019) 100151.
[28] G. Murtaza, L. Shuib, A.W. Abdul Wahab, G. Mujtaba, G. Mujtaba, H.F. Nweke, et al., Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges, Artif. Intell. Rev. 53 (2020) 1655–1720.
[29] J. Zhou, L.Y. Luo, Q. Dou, H. Chen, C. Chen, G.J. Li, et al., Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images, J. Magn. Reson. Imaging 50 (4) (2019) 1144–1151.
[30] H.D. Cheng, J. Shan, W. Ju, Y. Guo, L. Zhang, Automated breast cancer detection and classification using ultrasound images: a survey, Pattern Recognit. 43 (1) (2010) 299–317.
[31] F. Shahidi, S.M. Daud, H. Abas, N.A. Ahmad, N. Maarop, Breast cancer classification using deep learning approaches and histopathology image: a comparison study, IEEE Access 8 (2020) 187531–187552.
[32] M.Z. Alom, C. Yakopcic, M.S. Nasrin, T.M. Taha, V.K. Asari, Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network, J. Digit. Imag. 32 (2019) 605–617.
[33] S. Reis, P. Gazinska, J.H. Hipwell, T. Mertzanidou, K. Naidoo, N. Williams, et al., Automated classification of breast cancer stroma maturity from histological images, IEEE Trans. Biomed. Eng. 64 (10) (2017) 2344–2352.
[34] N.I. Yassin, S. Omran, E.M. El Houby, H. Allam, Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: a systematic review, Comput. Methods Programs Biomed. 156 (2018) 25–45.
[35] Y.M. George, H.H. Zayed, M.I. Roushdy, B.M. Elbagoury, Remote computer-aided breast cancer detection and diagnosis system based on cytological images, IEEE Syst. J. 8 (3) (2013) 949–964.
[36] S. Sharma, R. Mehra, Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images-a comparative insight, J. Digit. Imag. 33 (2020) 632–654.
[37] G. Chugh, S. Kumar, N. Singh, Survey on machine learning and deep learning applications in breast cancer diagnosis, Cogn. Comput. (2021) 1–20.
[38] G. Hamed, M.A.E.R. Marey, S.E.S. Amin, M.F. Tolba, Deep learning in breast cancer detection and classification, in: Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020), Springer International Publishing, 2020, pp. 322–333.
[39] M. Tiwari, R. Bharuka, P. Shah, R. Lokare, Breast cancer prediction using deep learning and machine learning techniques, 2020, Available at SSRN 3558786.
[40] A.U.R. Butt, W. Ahmad, R. Ashraf, M. Asif, S.A. Cheema, Computer aided diagnosis (CAD) for segmentation and classification of burnt human skin, in: 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), IEEE, 2019, July, pp. 1–5.
[41] D.A. Ragab, O. Attallah, M. Sharkas, J. Ren, S. Marshall, A framework for breast cancer classification using multi-DCNNs, Comput. Biol. Med. 131 (2021) 104245.
[42] R. Ashraf, I. Kiran, T. Mahmood, A.U.R. Butt, N. Razzaq, Z. Farooq, An efficient technique for skin cancer classification using deep learning, in: 2020 IEEE 23rd International Multitopic Conference (INMIC), IEEE, 2020, November, pp. 1–5.
[43] S. Khan, N. Islam, Z. Jan, I.U. Din, J.J.C. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, Pattern Recognit. Lett. 125 (2019) 1–6.
[44] F.A. Khan, A.U.R. Butt, M. Asif, W. Ahmad, M. Nawaz, M. Jamjoom, E. Alabdulkreem, Computer-aided diagnosis for burnt skin images using deep convolutional neural network, Multimed. Tools Appl. 79 (2020) 34545–34568.
[45] F.A. Khan, A.U.R. Butt, M. Asif, H. Aljuaid, A. Adnan, S. Shaheen, Burnt human skin segmentation and depth classification using deep convolutional neural network (DCNN), J. Med. Imaging Health Inform. 10 (10) (2020) 2421–2429.
[46] Z. Hameed, S. Zahia, B. Garcia-Zapirain, J. Javier Aguirre, A. Maria Vanegas, Breast cancer histopathology image classification using an ensemble of deep learning models, Sensors 20 (16) (2020) 4373.
[47] P. Gupta, S. Garg, Breast cancer prediction using varying parameters of machine learning models, Proc. Comput. Sci. 171 (2020) 593–601.
[48] X. Zhang, Y. Zhang, Q. Zhang, Y. Ren, T. Qiu, J. Ma, Q. Sun, Extracting comprehensive clinical information for breast cancer using deep learning methods, Int. J. Med. Inform. 132 (2019) 103985.
[49] J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, J. Fan, 2020, Deep learning assisted efficient AdaBoost algorithm for breast cancer detection.
[50] R. Krithiga, P. Geetha, Deep learning based breast cancer detection and classification using fuzzy merging techniques, Mach. Vis. Appl. 31 (2020) 1–18.
[51] Z. Zhang, E. Sejdić, Radiological images and machine learning: trends, perspectives, and prospects, Comput. Biol. Med. 108 (2019) 354–370.
[52] T. Pang, J.H.D. Wong, W.L. Ng, C.S. Chan, Deep learning radiomics in breast cancer with different modalities: overview and future, Expert Syst. Appl. 158 (2020) 113501.
[53] E.H. Houssein, M.M. Emam, A.A. Ali, P.N. Suganthan, Deep and machine learning techniques for medical imaging-based breast cancer: a comprehensive review, Expert Syst. Appl. 167 (2021) 114161.
[54] D. Singh, A.K. Singh, Role of image thermography in early breast cancer detection-past, present and future, Comput. Methods Programs Biomed. 183 (2020) 105074.
[55] G. Choy, O. Khalilzadeh, M. Michalski, S. Do, A.E. Samir, O.S. Pianykh, et al., Current applications and future impact of machine learning in radiology, Radiology 288 (2) (2018) 318–328.
[56] S.A. Medjahed, T.A. Saadi, A. Benyettou, Breast cancer diagnosis by using k-nearest neighbor with different distances and classification rules, Int. J. Comput. Appl. 62 (1) (2013).
[57] S. Joshi, B. Pandey, N. Joshi, Comparative analysis of Naïve Bayes and j48 classification algorithms, Int. J. Adv. Res. Comput. Sci. Softw. Eng. 5 (12) (2015) 813–817.
[58] S.B. Sakri, N.B.A. Rashid, Z.M. Zain, Particle swarm optimization feature selection for breast cancer recurrence prediction, IEEE Access 6 (2018) 29637–29647.
[59] M. Karabatak, A new classifier for breast cancer detection based on Naïve Bayesian, Measurement 72 (2015) 32–36.
[60] V. Chaurasia, S. Pal, Stacking-based ensemble framework and feature selection technique for the detection of breast cancer, SN Comput. Sci. 2 (2021) 1–13.
[61] S. Ibrahim, S. Nazir, S.A. Velastin, Feature selection using correlation analysis and principal component analysis for accurate breast cancer diagnosis, J. Imaging 7 (11) (2021) 225.
[62] T.E. Mathew, K.A. Kumar, A logistic regression based hybrid model for breast cancer classification, Indian J. Comput. Sci. Eng. 11 (6) (2020) 899–903.
[63] R. Sheikhpour, M.A. Sarram, R. Sheikhpour, Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer, Appl. Soft Comput. 40 (2016) 113–131.
[64] T.A. Assegie, R.L. Tulasi, N.K. Kumar, Breast cancer prediction model with decision tree and adaptive boosting, IAES Int. J. Artif. Intell. 10 (1) (2021) 184.
[65] T.S. Lim, K.G. Tay, A. Huong, X.Y. Lim, Breast cancer diagnosis system using hybrid support vector machine-artificial neural network, Int. J. Electr. Comput. Eng. (IJECE) 11 (4) (2021) 3059.
[66] P. Hamsagayathri, P. Sampath, Performance analysis of breast cancer classification using decision tree classifiers, Int. J. Curr. Pharm. Res. 9 (2) (2017) 19–25.
[67] A.E. Hassanien, Classification and feature selection of breast cancer data based on decision tree algorithm, Stud. Inform. Control 12 (1) (2003) 33–40.
[68] A.B. Yusuf, R.M. Dima, S.K. Aina, Optimized breast cancer classification using feature selection and outliers detection, J. Niger. Soc. Phys. Sci. (2021) 298–307.
[69] F. Abdat, M. Amouroux, Y. Guermeur, W.C.P.M. Blondel, Hybrid feature selection and SVM-based classification for mouse skin precancerous stages diagnosis from bimodal spectroscopy, Opt. Express 20 (1) (2012) 228–244.
[70] I. Vidić, L. Egnell, N.P. Jerome, J.R. Teruel, T.E. Sjøbakk, A. Østlie, et al., Support vector machine for breast cancer classification using diffusion-weighted MRI histogram features: preliminary study, J. Magn. Reson. Imaging 47 (5) (2018) 1205–1216.
[71] A.S. Elkorany, M. Marey, K.M. Almustafa, Z.F. Elsharkawy, Breast cancer diagnosis using support vector machines optimized by whale optimization and dragonfly algorithms, IEEE Access 10 (2022) 69688–69699.
[72] F.Z. Nakach, H. Zerouaoui, A. Idri, Random forest based deep hybrid architecture for histopathological breast cancer images classification, in: International Conference on Computational Science and Its Applications, Springer International Publishing, Cham, 2022, July, pp. 3–18.
[73] Y. Buttan, A. Chaudhary, K. Saxena, An improved model for breast cancer classification using random forest with grid search method, in: Proceedings of Second International Conference on Smart Energy and Communication, ICSEC 2020, Springer, Singapore, 2021, pp. 407–415.
[74] E. Akkur, F. Turk, O. Erogul, Breast cancer diagnosis using feature selection approaches and Bayesian optimization, Comput. Syst. Sci. Eng. 45 (2) (2023).
[75] M.H. Memon, J.P. Li, A.U. Haq, M.H. Memon, W. Zhou, Breast cancer detection in the IOT health environment using modified recursive feature selection, Wirel. Commun. Mob. Comput. 2019 (2019) 1–19.

[76] S. Salvi, A. Kadam, Breast Cancer Detection Using Deep Learning and IoT Technologies, J. Phys. Conf. Ser. 1831 (1) (2021, March) 012030, IOP Publishing.

[77] V.N. Gopal, F. Al-Turjman, R. Kumar, L. Anand, M. Rajesh, Feature selection and classification in breast cancer prediction using IoT and machine learning, Measurement 178 (2021) 109442.

[78] M. Lamba, G. Munjal, Y. Gigras, Supervising healthcare schemes using machine learning in breast Cancer and internet of things (SHSMLIoT), in: Internet of Healthcare Things: Machine Learning for Security and Privacy, 2022, pp. 241–263.

[79] P. Gupta, M. Ghosh, Revolutionizing healthcare with 5G, Telecom Bus. Rev. 12 (2019) 41.

[80] A. Mashekova, Y. Zhao, E.Y. Ng, V. Zarikas, S.C. Fok, O. Mukhmetov, Early detection of the breast cancer using infrared technology-a comprehensive review, Therm. Sci. Eng. Prog. 27 (2022) 101142.

[81] E. Selem, M. Fatehy, S.M. Abd El-Kader, E-health applications over 5G networks: challenges and state of the art, in: 2019 6th International Conference on Advanced Control Circuits and Systems (ACCS) & 2019 5th International Conference on New Paradigms in Electronics & Information Technology (PEIT), IEEE, 2019, November, pp. 111–118.

[82] S.A. Alanazi, M.M. Kamruzzaman, M.N. Islam Sarker, M. Alruwaili, Y. Alhwaiti, N. Alshammari, M.H. Siddiqi, Boosting breast cancer detection using convolutional neural network, J. Healthc. Eng. 2021 (2021).

[83] T. He, Y.Y. Pu, Y.Q. Zhang, Z.B. Qian, L.H. Guo, L.P. Sun, et al., 5G-based telerobotic ultrasound system improves access to breast examination in rural and remote areas: a prospective and two-scenario study, Diagnostics 13 (3) (2023) 362.

[84] M. Cabanillas-Carbonell, J. Pérez-Martinez, J.A. Yanez, 5G technology in the digital transformation of healthcare, a systematic review, Sustainability 15 (4) (2023) 3178.

[85] D. Saraswathi, E. Srinivasan, Performance analysis of mammogram CAD system using SVM and KNN classifier, in: 2017 International Conference on Inventive Systems and Control (ICISC), IEEE, 2017, January, pp. 1–5.

[86] D.E. Gbenga, N. Christopher, D.C. Yetunde, N. Maiduguri, Performance comparison of machine learning techniques for breast cancer detection, Nova J. Eng. Appl. Sci. 6 (1) (2017) 1–8.

[87] M. Amrane, S. Oukid, I. Gagaoua, T. Ensari, Breast cancer classification using machine learning, in: 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), IEEE, 2018, April, pp. 1–4.

[88] B.A. Mohamed, N.M. Salem, Automatic classification of masses from digital mammograms, in: 2018 35th National Radio Science Conference (NRSC), IEEE, 2018, March, pp. 495–502.

[89] T.T. Htay, S.S. Maung, Early stage breast cancer detection system using glcm feature extraction and k-nearest neighbor (k-NN) on mammography image, in: 2018 18th International Symposium on Communications and Information Technologies (ISCIT), IEEE, 2018, September, pp. 171–175.

[90] A. Bharat, N. Pooja, R.A. Reddy, Using machine learning algorithms for breast cancer risk prediction and diagnosis, in: 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C), IEEE, 2018, October, pp. 1–4.

[91] R.A.N. Diaz, N.N.T. Swandewi, K.D.P. Novianti, Malignancy determination breast cancer based on mammogram image with k-nearest neighbor, in: 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS), Vol. 1, IEEE, 2019, August, pp. 233–237.

[92] S. Mostafa, R. Mubarak, M. El-Adawy, A.F. Ibrahim, M.M. Gomaa, R.M. Kamal, Breast cancer detection using polynomial fitting applied on contrast enhanced spectral mammography, in: 2019 International Conference on Innovative Trends in Computer Engineering (ITCE), IEEE, 2019, February, pp. 11–16.

[93] R. MurtiRawat, S. Panchal, V.K. Singh, Y. Panchal, Breast Cancer detection using K-nearest neighbors, logistic regression and ensemble learning, in: 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), IEEE, 2020, July, pp. 534–540.

[94] S. Laghmati, B. Cherradi, A. Tmiri, O. Daanouni, S. Hamida, Classification of patients with breast cancer using neighbourhood component analysis and supervised machine learning techniques, in: 2020 3rd International Conference on Advanced Communication Technologies and Networking (CommNet), IEEE, 2020, September, pp. 1–6.

[95] A. Ivaturi, A. Singh, B. Gunanvitha, K.S. Chethan, Soft classification techniques for breast cancer detection and classification, in: 2020 International Conference on Intelligent Engineering and Management (ICIEM), IEEE, 2020, June, pp. 437–442.

[96] S. Chattopadhyay, A. Dey, P.K. Singh, R. Sarkar, DRDA-Net: dense residual dual-shuffle attention network for breast cancer classification using histopathological images, Comput. Biol. Med. 145 (2022) 105437.

[97] C.C. Ukwuoma, M.A. Hossain, J.K. Jackson, G.U. Nneji, H.N. Monday, Z. Qin, Multi-classification of breast cancer lesions in histopathological images using DEEP_Pachi: multiple self-attention head, Diagnostics 12 (5) (2022) 1152.

[98] M. Sharma, A. Mandloi, M. Bhattacharya, A novel DeepML framework for multi-classification of breast cancer based on transfer learning, Int. J. Imaging Syst. Technol. 32 (6) (2022) 1963–1977.