International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 6, December 2022, pp. 6397~6409
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i6.pp6397-6409

Breast cancer diagnosis: a survey of pre-processing, segmentation, feature extraction and classification

Ehsan Sadeghi Pour, Mahdi Esmaeili, Morteza Romoozi
Department of Electrical and Computer Engineering, Kashan Branch, Islamic Azad University, Kashan, Iran

Article Info

Article history:

Received Oct 13, 2021
Revised Jun 4, 2022
Accepted Jul 1, 2022

Keywords:

Breast cancer classification
Breast cancer diagnosis
Deep learning
Machine learning
Mammography

ABSTRACT

Machine learning methods have been an interesting method in the field of medical for many years, and they have achieved successful results in various fields of medical science. This paper examines the effects of using machine learning algorithms in the diagnosis and classification of breast cancer from mammography imaging data. Cancer diagnosis is the identification of images as cancer or non-cancer, and this involves image preprocessing, feature extraction, classification, and performance analysis. This article studied 93 different references mentioned in the previous years in the field of processing and tries to find an effective way to diagnose and classify breast cancer. Based on the results of this research, it can be concluded that most of today's successful methods focus on the use of deep learning methods. Finding a new method requires an overview of existing methods in the field of deep learning methods in order to make a comparison and case study.

Corresponding Author:

Mahdi Esmaeili
Department of Electrical and Computer Engineering, Kashan Branch, Islamic Azad University
Ostadan st., Qotb-e Ravandi st., Kashan, Isfahan, Iran
Email: M.esmaeili@iaukashan.ac.ir

1. INTRODUCTION

Breast cancer is one of the most serious and prevalent types of cancer, affecting the majority of women and being the main cause of death. Between 1 in 8 and 1 in 12 women in today's developed world will acquire breast cancer during their lifetime [1]. It is critical to forecast breast cancer risk in order to combat the disease. Breast cancer risk is classified into two categories [1]. The first type implies that a person may get breast cancer within a specified time period [1]. The second kind implies the probability of a high-risk gene mutation [2]. Breast tumors are an abnormal growth of breast tissue that may manifest as discharge from a lump or nipple or a change in the texture of the skin around the nipple area. Cancers are uncontrolled cell divisions that are capable of invading other tissues. Through the blood and lymphatic systems, cancer cells can move to different areas of the body [3]. This is the greatest cause of death among women in their forties and fifties [3]. Breast cancer is the second most frequent type of cancer and the main cause of cancer death in women, trailing only lung cancer [3]. In the previous 50 years, the disease has grown in prominence, and its prevalence has increased in recent years. Breast cancer screening guidelines are scarce at the moment. The Iranian preventive services working group recommends screening for women between the ages of 45 and 70 but provides no definitive recommendations for women older than this age range. Breast cancer risk prediction can be used to both incentivize high-risk women who are not already screened and ensure that screening criteria are followed by those who would not do so otherwise. A statistical model that predicts the risk of breast cancer can also be utilized to build cancer preventive and risk reduction measures [4]. As such, the goal of this project is to construct models capable of properly forecasting women's chance of developing breast cancer over the next five years. As a result, it is vital to discern between an accurate diagnosis and a correct decision when viewing mammography images.

Numerous previous studies have used the Gail model to estimate the risk of breast cancer. The Gail model is a statistical model that estimates the risk of breast cancer in women with no personal history of the disease and no known mutations in breast cancer-risk genes [5]. Historically, this model required six inputs: current age, menarche age, birth age, the number of first-degree relatives diagnosed with breast cancer, race/ethnic origin, and the number of previous breast biopsies [5]. The Gail model combines both generic and specific breast cancer risk rates, and these inputs (weighted using logistic regression) are used to anticipate a woman's likelihood of developing breast cancer [6]. Breast cancer risk assessment tool (BCRAT) is Gail model implementations that, when available, will take into account both the six classic Gail model inputs and atypical hyperplasia personal data history [7]. On the other hand, the Gail model is far from ideal. This strategy has been demonstrated to be unsuccessful in certain parts of the population for forecasting breast cancer risk [8], [9]. Furthermore, this method has difficulty differentiating breast cancer from non-breast cancer on an individual basis [10]. According to one study, logistic regression with a single age, previous breast biopsies, and family history of breast cancer as inputs successfully predicted the probability of five-year breast cancer [9].

Numerous earlier researches indicate how it is possible to improve the prediction power of the Gail model by utilizing expensive inputs. Earlier research has established that when evaluated using the same data set, models that incorporate one or more instructional inputs in addition to the Gail model predict better breast cancer risks than the Gail model. These works make use of simple statistical models and incorporate data gathered by costly and/or aggressive methods. Breast density measurements [11]–[13] genetic single-nucleotide polymorphism measurements [12], [14]–[16] nipple aspirate fluid cytology measurements [17], and/or hormone levels measurements [12], [18]. All of this was accomplished by employing simple statistical models such as Cox proportional hazards regressions [11], [17], logistic regressions [12], [16], or the Gail model. Due to these approaches' limitations in terms of high-precision diagnosis, machine learning-based methods have been created, which is the core emphasis of this work.

The study analyzes if modern machine learning algorithms can anticipate the risk of breast cancer more accurately than BCRAT can in the next five years. Complex machine learning models can exploit subtle patterns in the input data to improve forecasting accuracy. In [19] proved the efficacy of machine learning algorithms in predicting endometrial cancer risk using personal health data. In [19], six machine learning techniques were used to assess the probability of getting endometrial cancer: logistic regression, naive Bayesian, decision tree (DT), linear discriminant analysis (LDA), support vector machine (SVM), and artificial neural network (ANN). Internal validation was performed in the prostate, lung, colorectal, and ovarian cancer screening trials (PLCO) by training models on 70% of the data and evaluating their performance on the remaining 30% of the data [20]. A neural network was discovered to be the best predictor of risk in [19]. The goal of this study is to determine whether breast cancer risk can be predicted using machine learning algorithms, as well as endometrial cancer risk. By comparing the performance of alternative machine learning models to the BCRAT function, one can determine whether a model with a statistical foundation is more robust than BCRAT when given the same inputs and assessed on the same data set [7]. In [21], The goal of this study is to determine Lifestyles such as smoking and physical activity were suspected to affect breast cancer indirectly; however, comparing breast cancer patients between healthy and unhealthy lifestyles needs to be proven future. Lifestyles such as smoking and physical activity were suspected to affect breast cancer indirectly; however, comparing breast cancer patients between healthy and unhealthy lifestyles needs to be proven future. Changes lifestyle such as quitting smoking, active exercise, reducing alcohol consumption, taking vitamin and minerals seem an effective, easy, and economical ways to help prevention of breast cancer. In [22], this study aimed to investigate the knowledge, attitude and practice of women about breast cancer's screening methods in order to offer more appropriate training programs if necessary. A cross-sectional study was carried out with a population comprised of women who had referred to public health centers in Sanandaj in 2008. The results of this study do provide some understanding on the topic and suggest that although the majority of Iranian women seem to be quite knowledgeable about breast cancer and screening methods.

The remainder of this article is: the sections on pre-processing, segmentation and feature extraction, classification, and performance appraisal cover the stages involved in developing a computer-aided design model using machine learning methods. Each section explains the designated stage in detail and the machine learning approach used to perform it. The purpose of this article is to address the impact of applying machine learning principles and algorithms to histology and the research gaps associated with their implementation in a real-world setting. Finally, the conclusion makes recommendations for future research. The study's emphasis is on research employing mammography and histopathological pictures in all locations.

2. BREAST CANCER

A type of breast cancer that starts in women has symptoms such as a lump in the breast, a change in the shape of the breast, a dimple in the skin, discharge from the nipple, or scaling of part of the skin. In order to develop cancer, the gene for cell growth and proliferation must be altered. These mutations will then become a mass through cell proliferation. By identifying the gene that transmits this cancer, an important step can be taken in predicting breast cancer. One of the most important problems in showing the structure and macro performance of biological molecules is the high volume of genetic information. Also, one of the most important challenges in bioinformatics is that it requires the design and production of methods, algorithms, and tools to transform this vast amount of data, often heterogeneous (at low levels), into higher-level biological knowledge.

This article is based on breast cancer diagnosis and classification by considering mammography and histopathology images mammograms are X-ray scans of the breast that can detect symptoms of breast cancer in their early stages. A mammogram is produced using one of two ways. Digital mammography produces digital images, whereas film-screen mammography produces photographic film. Both techniques employ the same picture capture procedure. The woman undergoing the mammography will insert her breast between two transparent plates, which will squeeze it together and secure it. This helps to flatten the breast and prevents the image from becoming blurry. Two images of the breast are taken by the machine. The mammogram is then analyzed by a professional for anything abnormal that could be a sign of cancer. Any spot that does not appear to be normal tissue should be regarded as suspicious. The radiologist will search for regions of white, dense tissue and take note of its size, shape, and borders. On a mammography, a lump or tumor appears as a narrow white spot. Cancerous or benign tumors are also possible. A benign tumor does not pose a risk to health and is unlikely to grow or change shape. The majority of breast tumors are benign. Generally, small white flecks are benign. Their shape and pattern will be examined by the radiologist, as they can occasionally be indicative of cancer. A radiologist will search for anything odd on a mammography, in addition to thick breast tissue and suspected malignancies. Among the other anomalies: i) cysts, which are tiny sacs filled with fluid. The majority are benign cysts with a thin wall that are not malignant. If a doctor is unable to classify a cyst as a simple cyst, they may order additional testing to rule out cancer, ii) calcifications are calcium deposits. Macrocalcifications are larger deposits of calcium that typically arise as a result of aging. Microcalcifications are smaller deposits. Depending on the appearance of the microcalcifications, a doctor may do a test to see if they are indicative of malignancy, iii) fibroadenomas are benign breast tumors. They are spherical and may have the texture of a marble. Although fibroadenomas are more common in people in their 20 s and 30 s, they can occur at any age, and iv) scar tissue is frequently seen as white on mammograms. It is best to inform a doctor in advance of any scars on the breasts.

A mass may refer to a tumor, cyst, or fibroadenoma, regardless of whether they are cancerous. Additionally, a mammography might provide information regarding a person’s breast density. Individuals with thick breasts have a slightly increased risk of developing breast cancer. Dense breasts might make problems more difficult to detect on a mammogram. Mammograms are still possible after breast cancer surgery or implant placement. However, additional photos of each breast may be required, and image verification may take longer. Frequently, a radiologist will compare a mammogram to earlier pictures. This enables them to detect changes and determine whether an atypical area is indicative of malignancy.

Part of the X-ray radiation during the mammogram is absorbed in accordance with the tissue conditions, and the other part passes. Tissue in proportion to its nature absorbs some of the energy. The rate at which the signal leaves the cancerous tissue varies with the breast tissue. From the drop in the input to output signal, it can be determined whether the tissue has a cancerous mass. Today’s mammography is based on intelligent medical diagnostic systems, the main basis of which is image processing along with machine learning principles. Principles of image processing in intelligent medical systems are important for diagnosing breast cancer, because mammography images are inherently noisy, and this noise can be difficult to diagnose by a doctor. Of course, intelligent medical diagnostic systems have the ability to eliminate noise as well as diagnosis, but the doctor must give a definite opinion. Therefore, it is important to provide a smart medical diagnosis system to diagnose breast cancer. The use of image processing principles and techniques, along with statistical and cognitive identification of patterns in the diagnosis and automatic determination of breast cancer from mammography images has reduced human error and increased detection speed. In this study, we try to use the principles of image processing and machine learning to provide a system for diagnosing breast cancer tumors and separating benign and malignant conditions.

3. PRE-PROCESSING

Pre-processing is essential to reduce the complexity and efficiency of the image. The pre-processing stage minimizes image noise and aids in identifying focal points [23]. To boost local contrast, a limited contrast-compatible histogram alignment is applied [24]. Thresholds are used to minimize image noise. As a threshold value on the date of the picture intensity diagram, the following pixels are considered noise [23]. The Otsu threshold method is used to determine the ideal threshold [25]. To improve threshold results, background correction and filtering are applied. Fixed field correction lessens the effect of changing lighting conditions, while the filter reduces image noise [24]. The preceding strategies are widely used to enhance the quality of digital pictures used in machine learning applications. Data amplification and color normalization are the most often utilized pre-processing processes for computer-aided learning systems based on deep learning. Scaling, flipping, mirroring, blurring, and adding noise to training data sets are all examples of data augmentation. These changes correct the morphology of the image [26]. Additionally, histopathological images vary in hue and brightness due to changes in histopathology slide preparation (stain staining) and various noise circumstances when digital images are taken. As a result, stain normalization is a critical step in the pre-processing of histopathological imaging [26]. A whole-colored histopathology picture is used as the reference image in this procedure. The color values are adjusted to match the target image's color values. According to a study published in [27], stain normalization improves the performance of deep-learning classifiers. Stain normalization was employed by researchers in [28], [29] to develop stain-based computer-aided design tools for breast cancer histopathology. The study [30] presents a new evaluation of color staining normalization approaches for histopathology pictures. Certain research investigations have demonstrated extremely successful methods for normalizing color in digital photographs. For instance, Reinhard *et al.* [31] suggested a method for converting an red green blue (RGB) image to a color space based on perception (lab). By computing the mean and standard deviation for each axis of the image, the approach transfers color between the source and target images. In [32], researchers suggested a quantitative analysis-based strategy for adjusting the color of histopathology pictures. Khan *et al.* [33] proposed a staining method in which the image stain matrix is calculated using a stain color descriptor. Following that, the image's varying stain concentration values are retrieved using color decomposition (stain separation) and passed to the nonlinear mapping function (based on spline). Finally, the source image is recreated utilizing stained channels in the normal manner. Stain normalization should be used in conjunction with color enhancement to boost the performance of deep-learning computer-aided design systems, according to a recent study [26].

In [34], a review study and comparison of noise reduction methods in mammography images has been performed and the strengths and weaknesses of these methods in the pre-processing phase have been discussed. Also in [35], an analytical method of noise is presented in X-ray mammography images based on non-local mean method. In [36], a comparative adaptive weighted frost filter has been used for pre-processing to reduce noise in mammography images. In [37], the impact noise reduction in ultrasound images was performed using the modified Bayes method. This study is based on a combination of intelligent systems in [38]. The findings demonstrated the accuracy of breast cancer prediction and diagnosis. Although adaptive fuzzy neural networks were used in this study for adaptation and learning, given the difficulty associated with teaching this type of network, a combination of evolutionary algorithms and data mining systems may be a novel idea for increasing the efficiency of prediction and estimation of adaptive fuzzy neural networks with greater accuracy.

4. SEGMENTATION AND FEATURE EXTRACTION

After the pre-processing stage, they become vectors of a particular property [39]. Extraction plays a vital role in the design of machine learning models [40]. Extraction techniques for features computer-aided design can be divided into two categories: features that are handcrafted (manual design) and features that are learned using deep learning. When constructing handcrafted features, the user must exert some effort to determine whether characteristics are diagnostically acceptable. In contrast, deep learning models, while teaching networking for classification, automatically determine key features. These two sections are explained separately and an overview of its methods is given.

4.1. Handcrafted-based feature extraction methods

A computer-aided design system based on the division of graphical diagrams with space-color, extraction of tissue properties (such as gray-level co-occurrence matrix (GLCM), grassland-based ruminant livestock models (GRLM) and Euler methods) and classification based on linear discriminative analysis is presented in [41]. The system categorized 70 histopathology and mammography images with 100% accuracy. In [42], an accuracy of 97.75% was achieved using a feature-oriented vocabulary learning algorithm on a data set of human intraocular lesions and animal diagnostic laboratories. A computer-aided design method based on the extraction of morphological features was shown to be 85.7% accurate in classifying histopathological and mammographic data sets from 70 images [43]. However, this research has a key disadvantage in that it examined limited unpublished data sets for studies. Given this constraint, it is difficult to compare these efforts to other research. To circumvent this constraint, the authors made publicly available a set of histopathology data and breast cancer mammography images dubbed BreakHis data and conducted preliminary trials utilizing manual features. The authors discovered that parameter-free threshold adjacency statistics (PFTAS) has the highest cancer detection rates in comparison to local binary pattern (LBP), completed local binary patterns (CLBP), local phase quantization (LPQ), opening range breakout (ORB), and GLCM. Another study used the Gabor filter, wavelet, and LBP features in conjunction with an ensemble classifier to obtain 200x magnification and 90.32% accuracy for BreakHis picture classification. PFTAS features, combined with a non-parametric classification method based on multi-sample learning, recently achieved the greatest patient recognition rate (PRR) for the BreakHis data set [44]. However, the accuracy and standard deviation of this research are quite high.

In [45], the use of cross-section and extraction of edge-based and area-based features in a four-step system is ensured on a mammographic scale. Also in [46], wavelet transform and genetic algorithm are used to segment and extract the characteristics of mammography images. The use of fuzzy logic for region of interest is also provided for this purpose in [47]. In [48] used adaptive local thresholding methods and improved morphology to extract and segment mammography and histochemistry images. In [49], a cellular neural network was used along with region growing with the aim of segmentation and extraction of mammographic image characteristics. The use of micro array images to diagnose breast cancer masses has also been studied [50]. The use of back propagation neural networks to dissect and diagnose breast cancer masses has also been discussed in [51]. The application of business theory classification method based on Bayes theory in mammography images is also presented in [52]. In [53], a comparatively intelligent decision-making system has been used to diagnose breast cancer based on mammographic images based on evolutionary methods based on regression. In [54], the prediction of breast cancer recurrence is presented using optimized ensemble learning or hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR).

4.2. Deep learning-based feature extraction methods

Deep learning-based feature extraction methods is highly regarded by researchers for classifying histopathology and mammography images. In [55] represented that deep learning-based methods are more accurate than autoimmune colors, textures, and geometric features for automatic detection of carcinoma invasive tissues in images of total breast cancer slides. In [56], studies were conducted on histological data from breast cancer using the LBP, CLBP, LPQ, ORB, GLCM, and PFTAS characteristics. Subsequently, the same authors used a deep learning network to classify the BreakHis dataset in [57]. This model outperforms classification models based on LBP, CLBP, LPQ, ORB, and GLCM features in terms of accuracy. Recently, neural network convolution based on multivariate learning produced the maximum PRR for the BreakHis data set [44].

The use of the convolutional neural network for segmentation and feature extraction is presented in [58] along with a decoder and encryption mode. In [59], deep learning based on the conductive UNet 2 method has been used to segment breast tissue and fibroglandular. In [60], multi-task segmentation is presented in several sections of mammography images to find deep learning breast masses and standard convolutional neural network (CNN) methods. In [61], deep learning and the CNN method of V-net convolution have been used to segment mammography images as well as prostate images. Mammography imaging with the aim of diagnosing and classifying benign and malignant masses with an optimal area growth approach is presented in [62], which is based on dragonfly optimization algorithm and a combined approach of GLCM and GLRLM method for extracting features as input in the method and classification by the feed-forward neural network or feed forward neural network (FFNN) has been used with back propagation training. In [63], the proposed approach is used to crop the region of interests (ROIs) manually. Based on that numbers of features are extracted. In this proposed method a novel hybrid optimum feature selection (HOFS) method is used to find out the significant features to reach maximum accuracy for this classification. A number of selected features are applied to train the neural network. In this proposed method accessible informational index from the mini-mammographic image analysis society (MIAS) database was used. In [64], the purpose of this article was to review various approaches to detecting breast cancer using artificial intelligence (AI) and image processing. The authors present an innovative approach for identifying breast cancer using machine learning methods. Compared to current approaches, such as CNN, our particle swarm optimized wavelet neural network (PSOWNN) method appears to be relatively superior.

5. CLASSIFICATION

Classification is the process of categorizing data based on inherent quantitative information [65]. To diagnose breast cancer, data points should be classed as benign or malignant. Two stages comprise the categorization model: training and testing. The classifier is trained by providing it with features vectors containing class labels as input. These feature vectors can be thought of as learning instances for the classifier. After teaching samples to classify, they can be utilized to evaluate a feature vector corresponding to an unknown class. The output of the test is a class tag attached to a feature vector. The DT, the naive Bayesian technique, the K-nearest neighbor (KNN) approach, the ANN, SVM, and ensemble classifier are all examples of machine learning approaches that have been widely used to develop computer-aided design learning models [66], [67]. In [68] explored the applications, advantages, and limitations of these classifiers for cancer prediction and detection. This section discusses classification approaches that utilize deep learning systems and neural networks.

The use of ensemble classifiers has been fully studied in [68]. In this study, 193 articles published from 2000 to 2019 were compared with 9 different evaluation criteria in terms of efficiency. The use of histopathological and mammographic images with the aim of classifying cancer masses in the breast area has been reviewed in [69] based on machine learning methods. In [70], [71], the XCSR and XCSLA classification system for diagnosing some of the diseases are that it can also be used to diagnose breast cancer. In [72] another study, the use of an extreme learning machine with deep convolution features was presented with the aim of diagnosing and classifying breast cancer. In fact, the use of extreme learning machines is aimed at improving accuracy in the classification.

In [73], deep learning training for the classification of histopathological and mammography images for the cancer masses detection in breast area has been proposed. The cold metal transfer (CMT) data set was intended to be used in this study. Also, the deep neural network type was VGGNet-16, which achieved a high accuracy of about 93% to 97%. In [74], a combination of a support vector machine method with a deep neural network with the aim of classifying mammographic images to detect cancer masses has been used. The accuracy of this method was about 94% to 98% depending on the different data sets. In [75], [76], CNN has been used to diagnose and classify masses in the breast area, the results demonstrated by classical methods such as SVM, naïve Bayesian and other neural networks.

In [77], an end-to-end training approach with deep learning proposed for breast cancer classification by using curated breast imaging subset-digital database for screening mammography (CBIS-DDSM). In [78], one of the best review articles has been compiled which notice to the main points and challenges of deep learning for breast cancer classification in imaging data. Advances in deep learning for cancer diagnosis, particularly breast cancer, were discussed in [79]. Deep learning algorithms displayed expert-level performance in detecting breast cancer metastases in lymph nodes, outperforming prior feature-engineered methods of histopathology analysis. Additionally, this study on deep learning enables large-scale morphology-based research, as demonstrated recently in the mapping and analysis of tumor infiltrating lymphocyte patterns in hundreds of specimens from the Cancer Genome Atlas digital slide archive. Another systematic review [80] summarized the current state of the art for computer-aided diagnosis methods for breast cancer. Based on this research collecting data and processing to find an optimal solution of breast cancer diagnosis and classification proposed based on machine learning and as the main results, deep learning methods have the most advantages for this job. Another semantic segmentation and classification of breast cancer masses proposed in [81] which used deep learning. In this method, human epidermal growth factor receptor-2 deep neural network (Her2Net) and trapezoidal long short-term memory (TLSTM) used as deep learning algorithm to segment and classify cell membranes and nuclei in breast area for evaluation. This method had high accuracy about 98.33%. In this study [82], both of linear discriminant analysis and SVM are compared by looking from accuracy, sensitivity, specificity, and F1-score. We will know which methods are better in classifying breast cancer dataset. The result shows that the support vector machine has better performance than the linear discriminant analysis. It can be seen from the accuracy is 98.77%.

6. METHODS AND RESULTS

To this day, the following methods of breast tumor diagnosis in frequency images have been presented: i) morphological methods, ii) based on machine learning, iii) based on fuzzy logic, iv) based on evolutionary algorithms, and v) methods implementing chaos theory. These algorithms also include hybrid models based on the following single algorithms: i) C-means, ii) K-means, iii) fuzzy C-means (FCM), iv) fisher, and v) H-means. The hybrid diagnosis models are based on the following algorithms: i) genetic algorithm and FCM, ii) diagnosis based on multilayer perceptron neural network and optimized particle swarm algorithm. The following neural network-based classification methods are recommended: i) multilayer perceptron neural network, ii) adaptive neural oscillator network, iii) Hopfield neural network, iv) probabilistic neural network, v) radial basis function neural network, vi) self-organizing map neural network, vii) neocognitron neural network, and viii) Grossberg neural network. The other recommended evolutionary algorithms to improve segmentation and create optimal classes for breast cancer diagnosis from images include: i) gray wolf optimization algorithm, ii) dragonfly algorithm, iii) social spider algorithm, iv) bacterial foraging algorithm, v) ant colony algorithm, and vi) multi-objective genetic algorithm.

Adaptive filters such as least mean square (LMS), recursive least square (RLS) or normalized least mean square (NLMS) used in signal processing is a novel approach that combines image and signal processing to compress images into matrices with rows and columns of features. Then, a dimensionality reduction method should select and extract features to remove redundant features. The fractal theory is another similar compression method, and the Bézier or Brownian curves can also be used for this purpose. Chaos theory can use Chebyshev, Lyapunov, Lorenz, and other equations for stable and resistant segmentation, and despite their computational complexity, can greatly improve accuracy and sensitivity. Table 1 (see in appendix) summarizes the important breast tumor diagnosis and classification studies with the strengths and weaknesses of proposed methods. To compare the various methods presented so far in terms of evaluation criteria, most studies use accuracy percentage as the main breast cancer diagnosis and classification comparison criterion. Table 2 compares studies using the same dataset, e.g. the MIAS dataset. A comparison can be made here between classical methods, deep learning algorithms and extreme learning machines which are listed in Table 3.

In general, it can be argued that edge localized mode (ELM) is exactly the opposite of deep learning methods and alternative classification methods such as SVM and Naïve Bayesian. The ELM algorithm can employ nonlinear activation functions such as sigmoid or sinusoidal or non-derivative activation functions as well as using a linear function to activate cells or neurons in the hidden layer because of its high flexibility.

Table 2. Compares studies using the same dataset

Reference	Accuracy (%)
Rouhi *et al.* 2015 [49]	96.47 %
Khalilabad and Hassanpour 2017 [50]	95.45 %
Kaymak *et al*. 2017 [51]	70.40 %
Karabatak 2015 [52]	98.54 %
Wang *et al.* 2018 [53]	97.10 %
Geweid and Abdallah 2019 [83]	85 %

Table 3. A case study for breast cancer diagnosis and classification

Naïve Bayesian	SVM	CNN	Recursive neural network (RNN)	ELM
Slow training mechanism	Slow training mechanism	Slow training mechanism	Slow training mechanism	Slow training mechanism
binary classification ability	binary classification ability	Multi-class classification ability	Multi-class classification ability	Multi-class classification ability and also multi-objective and real-time
Quadratic Programming	Quadratic programming	Non-linear multi-objective programming	Multi-objective Quadratic programming	Non-linear multi-class and multi-objective programming
Improved evaluation criteria such as accuracy, sensitivity, and feature rate	Improved diagnosis and classification of benign and malignant tumors and determination of exact tumor area	Improved accuracy, features, and sensitivity relative to the New Bayesian method, random forest algorithm, support vector machine, and K nearest neighbor	Accurate tumor diagnosis and identification	Accurate tumor diagnosis and identification
High computational complexity and run time	Failure to discriminate and classify benign and malignant tumors and compare the proposed approach with previous deep learning methods	Very computationally and time intensive Slow tumor diagnosis	Very computationally intensive and inaccurate comparison without mentioning used data (methods should compare the data in a similar body state.)	Very computationally-intensive

7. CONCLUSION

According to studies in the field of breast cancer diagnosis with machine learning methods, it was observed that there are several basic steps in it, which include pre-processing, segmentation and features extraction, and finally classification. A variety of methods have been studied in these three sections from 93 different references over the years. Based on the available analysis and results of these articles, it can be seen deep learning methods have high capabilities in the field of pre-processing, segmentation and features extraction, and also in classification. However, a variety of studies have been performed with these methods, including the CNN and other similar methods combined with other methods. Therefore, it will be interesting to present an intelligent and automatic method that can detect and classify benign and malignant cancer masses from mammography, histopathology and any other type of visual data in the field of breast cancer diagnosis. Also, considering evaluation criteria such as accuracy, sensitivity, specificity, mean square error, recall, receiver operating characteristic (ROC), area under curve (AUC) rate and other evaluation criteria is necessary to compare with previous methods especially with deep learning methods. Therefore, in future research, a combined method will be used with the opposite point of deep learning algorithms, namely, extreme learning machine with combination to some optimal methods.

8. FUTURE WORKS

This study was the point of view of recent methods for pre-processing, segmentation with feature extraction and also classification of breast cancer. Based on this study, it is observed that deep learning methods obtained the best results in any parts of breast cancer classification and diagnosis and also in pre-processing, segmentation and feature extraction. Creating classes for determining benign and malignant masses in breast are I one of the open challenges. Another challenge is to find the exact area of tumor and also estimating its size. Also increasing some evaluation criteria such as accuracy, sensitivity, specificity, ROC, AUC, runtime, and others, is the main parts of these challenges in any parts. So, as the future works, it will be study about different parts of breast cancer diagnosis and classification: i) pre-processing by proposing optimized and enhanced method for reducing mammogram noises; ii) proposing evolutionary algorithm for segmentation and feature extraction to find the best features of images. This is a vital part of this algorithm due to finding exact area of masses in images which will be then classify to find the types of masses; and iii) classification which divided images and founded masses in three parts such as benign, malignant and suspicious. These three creativity and contribution will be explained in details in other articles.

APPENDIX

Table 1. Summary of reviewed methods *(continue)*

Reference	Method	Advantages	Disadvantages
Rouhi *et al.* 2015 [49]	Cellular neural network regional growth segmentation with a specific threshold and improved classification parameters based on a genetic algorithm in mammographic images	Improved accuracy, features, and sensitivity relative to the new Bayesian method, random forest algorithm, support vector machine, and K nearest neighbor	Very computationally and time intensive Slow tumor diagnosis
Khalilabad and Hassanpour 2016 [50]	Breast tumor diagnosis using microarray mammographic images	More accurate tumor diagnosis	Incorrect diagnosis of the tumor areas Noisy images after processing
Kaymak *et al.* 2017 [51]	Mammographic image classification for breast cancer diagnosis using the backpropagation neural network	Determination of tumor area	Poor tumor detection accuracy from datasets and slow and computationally-intensive processing
Karabatak 2015 [52]	Tumor classification and diagnosis using the new Bayesian method	Improved evaluation criteria such as accuracy, sensitivity, and feature rate	High computational complexity and run time
Wang *et al.* 2018 [53]	Adaptive intelligent decision-making system for breast cancer diagnosis from mammographic images using regression analysis	Ability to diagnose tumors and estimate life expectancy with or without tumors and their exact location in images	Poor tumor location accuracy with high computational complexity
Kaur *et al.* 2019 [74]	Segmentation-based breast tumors diagnosis and classification from mammographic images using the K-means and SURF algorithms and hybrid multiclass support vector machine-deep learning classification	Improved diagnosis and classification of benign and malignant tumors and determination of exact tumor area	Failure to discriminate and classify benign and malignant tumors and compare the proposed approach with previous deep learning methods
Yassin *et al.* 2018 [80]	Review Study	Analysis of intelligent breast cancer diagnosis methods and more efficient neural networks	Failure to evaluate denoising and segmentation methods before feature extraction and final classification
Geweid and Abdallah 2019 [83]	Breast cancer diagnosis using M-level optimization functions based on a non-parametric pixel intensity method	Detailed analysis of mammography images and malignant tumors with partial differential equation, non-coplanar motion in breast cancer and its overall dynamics, five simple differential equations, and estimation of intensity using nonparametric techniques are useful as comparison criteria.	High computational complexity and failure to consider suspicious and benign tumors and specify data Inaccurate comparison without stating used data (methods should compare the data in a fixed body state.)
Guo *et al.* 2017 [84]	Review study	Analysis of intelligent breast cancer diagnosis methods and more efficient neural networks	Failure to evaluate denoising and segmentation methods before feature extraction and final classification

Table 1. Summary of reviewed methods

Reference	Method	Advantages	Disadvantages
Patel and Sinha 2014 [85]	Breast cancer diagnosis from mammographic images using feature analysis, preprocessing, and an optimal classifier	Better evaluation criteria including accuracy, sensitivity, and feature rate and diagnosis of benign and malignant tumors	High computational complexity and run time
Singh and Gupta 2015 [86]	Breast cancer diagnosis from mammographic images with preprocessing followed by smoothing and thresholding for feature extraction Continuous window-finding with lower-variance min-max to determine tumor areas and sizes in images based on morphological operations and image gradient technique.	Improved evaluations criteria including accuracy, sensitivity, and feature rate with correct diagnosis of the tumor area	High computational complexity and run time
Tang *et al*. 2009 [87]	Review Study	Analysis of intelligent breast cancer diagnosis methods and more efficient neural networks	Failure to evaluate denoising and segmentation methods before feature extraction and final classification
Lee 2019 [88]	Performance analysis of the Compton camera with the Si/CZT lens for breast tumor diagnosis using the Monte Carlo method	Accurate determinations of tumor area in two and three-dimensional images	Inability to segment and classify benign and malignant tumors and analyze results and evaluation criteria
Khan *et al*. 2019 [89]	Implementing the GoogLeNet, VGGNet, and ResNet deep learning architectures to diagnose and classify breast cancer from mammographic images	Improved diagnosis and classification of benign and malignant tumors and determination of exact tumor area	Failure to mention the deep learning structure created in hidden layers, including fully connected layers, pooling layers, and the convolve layer, High computational complexity and run time in diagnosis and classification
Rahmatinia and Fahimi 2017 [90]	Breast tumor diagnosis with thermographic methods and high-frequency stimulation based on radiofrequency	Stimulation and accurate determination of tumor area	Computationally-intensive without discrimination of tumor states after diagnosis
Wang *et al*. 2019 [91]	Breast cancer diagnosis based on fusion features, convolution neural network, and extreme learning machine (ELM) classification	Accurate tumor diagnosis and identification	Very computationally-intensive
Li *et al*. 2019 [92]	Extracting a distinct pattern for breast cancer histopathological image classification using the automatic structure based on convulsion neural network and support vector machine	Accurate tumor diagnosis and identification	Very computationally intensive and inaccurate comparison without mentioning used data (methods should compare the data in a similar body state).
Panesar *et al*. 2017 [93]	Breast cancer diagnosis using a biosensor structure with quantum dots for tracking breast cancer mRNAs	Applying biosensor principles to diagnose breast tumors and present a novel optical and quantum processing method	Failure to review denoising and segmentation methods before feature extraction and final classification - Failure to present final results and evaluation and comparison

REFERENCES

[1] D. G. R. Evans and A. Howell, "Breast cancer risk-assessment models," *Breast Cancer Research*, vol. 9, no. 5, Oct. 2007, doi: 10.1186/bcr1750.

[2] J. M. Mehta *et al.*, "Breast cancer screening: women's attitudes and beliefs in light of updated united states preventive services task force and american cancer society guidelines," *Journal of Women's Health*, vol. 28, no. 3, pp. 302–313, Mar. 2019, doi: 10.1089/jwh.2017.6885.

[3] R. M. Pfeiffer *et al.*, "Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies," *PLoS Medicine*, vol. 10, no. 7, Jul. 2013, doi: 10.1371/journal.pmed.1001492.

[4] G. F. Stark, G. R. Hart, B. J. Nartowt, and J. Deng, "Predicting breast cancer risk using personal health data and machine learning models," *Plos One*, vol. 14, no. 12, Dec. 2019, doi: 10.1371/journal.pone.0226765.

[5] D. Sa-nguanraksa, T. Sasanakietkul, C. O-Charoenrat, A. Kulprom, and P. O-Charoenrat, "Gail model underestimates breast cancer risk in Thai population," *Asian Pacific Journal of Cancer Prevention*, vol. 20, no. 8, pp. 2385–2389, Aug. 2019, doi: 10.31557/APJCP.2019.20.8.2385.

[6] M. H. Gail *et al.*, "Projecting individualized probabilities of developing breast cancer for white females who are being examined annually," *JNCI Journal of the National Cancer Institute*, vol. 81, no. 24, pp. 1879–1886, Dec. 1989, doi: 10.1093/jnci/81.24.1879.

[7] C.-H. Lin *et al.*, "Contrasting epidemiology and clinicopathology of female breast cancer in Asians vs the US population," *JNCI: Journal of the National Cancer Institute*, vol. 111, no. 12, pp. 1298–1306, Dec. 2019, doi: 10.1093/jnci/djz090.

[8] J. G. Elmore and S. W. Fletcher, "The risk of cancer risk prediction: 'What Is my risk of getting breast cancer?,'" *JNCI: Journal of the National Cancer Institute*, vol. 98, no. 23, pp. 1673–1675, Dec. 2006, doi: 10.1093/jnci/djj501.

[9] R. T. Chlebowski *et al.*, "Predicting risk of breast cancer in postmenopausal women by hormone receptor status," *JNCI: Journal of the National Cancer Institute*, vol. 99, no. 22, pp. 1695–1705, Nov. 2007, doi: 10.1093/jnci/djm224.

[10] B. Rockhill, D. Spiegelman, C. Byrne, D. J. Hunter, and G. A. Colditz, "Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention," *JNCI Journal of the National Cancer Institute*, vol. 93, no. 5, pp. 358–366, Mar. 2001, doi: 10.1093/jnci/93.5.358.

[11] J. A. Tice, S. R. Cummings, R. Smith-Bindman, L. Ichikawa, W. E. Barlow, and K. Kerlikowske, "Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model," *Annals of Internal Medicine*, vol. 148, no. 5, Mar. 2008, doi: 10.7326/0003-4819-148-5-200803040-00004.

[12] X. Zhang *et al.*, "Addition of a polygenic risk score, mammographic density, and endogenous hormones to existing breast cancer risk prediction models: A nested case-control study," *Plos Medicine*, vol. 15, no. 9, Sep. 2018, doi: 10.1371/journal.pmed.1002644.

[13] H. Darabi, K. Czene, W. Zhao, J. Liu, P. Hall, and K. Humphreys, "Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement," *Breast Cancer Research*, vol. 14, no. 1, Feb. 2012, doi: 10.1186/bcr3110.

[14] M. E. Mealiffe, R. P. Stokowski, B. K. Rhees, R. L. Prentice, M. Pettinger, and D. A. Hinds, "Assessment of clinical validity of a breast cancer risk model combining genetic and clinical information," *JNCI Journal of the National Cancer Institute*, vol. 102, no. 21, pp. 1618–1627, Nov. 2010, doi: 10.1093/jnci/djq388.

[15] G. S. Dite *et al.*, "Using SNP genotypes to improve the discrimination of a simple breast cancer risk prediction model," *Breast Cancer Research and Treatment*, vol. 139, no. 3, pp. 887–896, Jun. 2013, doi: 10.1007/s10549-013-2610-2.

[16] S. Wacholder *et al.*, "Performance of common genetic variants in breast-cancer risk models," *New England Journal of Medicine*, vol. 362, no. 11, pp. 986–993, Mar. 2010, doi: 10.1056/NEJMoa0907727.

[17] J. A. Tice, R. Miike, K. Adduci, N. L. Petrakis, E. King, and M. R. Wrensch, "Nipple aspirate fluid cytology and the gail model for breast cancer risk assessment in a screening population," *Cancer Epidemiology, Biomarkers and Prevention*, vol. 14, no. 2, pp. 324–328, Feb. 2005, doi: 10.1158/1055-9965.EPI-04-0289.

[18] T. V Clendenen *et al.*, "Breast cancer risk prediction in women aged 35-50 years: impact of including sex hormone concentrations in the Gail model," *Breast Cancer Research*, vol. 21, no. 1, Dec. 2019, doi: 10.1186/s13058-019-1126-z.

[19] G. R. Hart, B. Nartowt, W. Muhammad, Y. Liang, G. Huang, and J. Deng, "Endometrial cancer risk prediction and stratification using personal health data," *International Journal of Radiation Oncology*Biology*Physics*, vol. 105, no. 1, Sep. 2019, doi: 10.1016/j.ijrobp.2019.06.2170.

[20] B. S. Kramer, J. Gohagan, P. C. Prorok, and C. Smart, "A national cancer institute sponsored screening trial for prostatic, lung, colorectal, and ovarian cancers," *Cancer*, vol. 71, no. S2, pp. 589–593, Aug. 2010, doi: 10.1002/cncr.2820710215.

[21] S. Solikhah, K. N. A. Setyawati, and M. Sangruangake, "Lifestyle breast cancer patients among Indonesian women: A nationwide survey," *International Journal of Public Health Science (IJPHS)*, vol. 10, no. 4, Dec. 2021, doi: 10.11591/ijphs.v10i4.20913.

[22] B. M. Behbahani, "Knowledge, attitude and practice of iranian women towards breast cancer screening methods," *International Journal of Public Health Science (IJPHS)*, vol. 6, no. 4, pp. 351–355, Dec. 2017, doi: 10.11591/ijphs.v6i4.10782.

[23] C. Demir and B. Yener, "Automated cancer diagnosis based on histopathological images : a systematic survey," *Rensselaer Polytechnic Institute, Tech. Rep*, 2005

[24] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems*, Elsevier, 1994, pp. 474–485, doi: 10.1016/B978-0-12-336156-1.50061-6.

[25] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[26] M. A. Aswathy and M. Jagannath, "Detection of breast cancer on digital histopathology images: Present status and future possibilities," *Informatics in Medicine Unlocked*, vol. 8, pp. 74–79, 2017, doi: 10.1016/j.imu.2016.11.001.

[27] F. Ciompi *et al.*, "The importance of stain normalization in colorectal tissue classification with convolutional networks," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Apr. 2017, pp. 160–163, doi: 10.1109/ISBI.2017.7950492.

[28] Z. Gandomkar, P. C. Brennan, and C. Mello-Thoms, "MuDeRN: Multi-category classification of breast histopathological image using deep residual networks," *Artificial Intelligence in Medicine*, vol. 88, pp. 14–24, Jun. 2018, doi: 10.1016/j.artmed.2018.04.005.

[29] D. M. Vo, N.-Q. Nguyen, and S.-W. Lee, "Classification of breast cancer histology images using incremental boosting convolution networks," *Information Sciences*, vol. 482, pp. 123–138, May 2019, doi: 10.1016/j.ins.2018.12.089.

[30] S. Roy, A. kumar Jain, S. Lal, and J. Kini, "A study about color normalization methods for histopathology images," *Micron*, vol. 114, pp. 42–61, Nov. 2018, doi: 10.1016/j.micron.2018.07.005.

[31] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Applications*, vol. 21, no. 4, pp. 34–41, 2001, doi: 10.1109/38.946629.

[32] M. Macenko *et al.*, "A method for normalizing histology slides for quantitative analysis," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Jun. 2009, pp. 1107–1110, doi: 10.1109/ISBI.2009.5193250.

[33] A. M. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1729–1738, Jun. 2014, doi: 10.1109/TBME.2014.2303294.

[34] S. Kannan, N. P. Subiramaniyam, A. T. Rajamanickam, and A. Balamurugan, "Performance comparison of noise reduction in mammogram images," *International Journal of Research in Engineering and Technology*, vol. 05, pp. 31–33, 2016.

[35] X. Jiang, Z. Wang, L. Zhang, and M. Stampanoni, "Noise-analysis-based non-local means method for x-ray grating-based mammography denoising," *IEEE Transactions on Nuclear Science*, vol. 60, no. 2, pp. 802–809, Apr. 2013, doi: 10.1109/TNS.2013.2240314.

[36] M. Talha, G. Bin Sulong, and A. Jaffar, "Preprocessing digital breast mammograms using adaptive weighted frost filter," *Biomedical Research*, vol. 27, no. 4, pp. 1407–1412, 2016.

[37] I. Elyasi, M. A. Pourmina, and M.-S. Moin, "Speckle reduction in breast cancer ultrasound images by using homogeneity modified bayes shrink," *Measurement*, vol. 91, pp. 55–65, Sep. 2016, doi: 10.1016/j.measurement.2016.05.025.

[38] E. Sadeghipour, N. Sahragard, M.-R. Sayebani, and R. Mahdizadeh, "Breast cancer detection based on a hybrid approach of firefly algorithm and intelligent systems," *Indian Journal of Fundamental and Applied Life Sciences*, vol. 5, pp. 468–472, 2015.

[39] S. B. Kotsiantis, "Supervised machine learning: a review of classification techniques," *Informatica (Slovenia)*, vol. 31, pp. 249–268, 2007

[40] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *2014 Fourth International Conference on Advanced Computing and Communication Technologies*, Feb. 2014, pp. 5–12, doi: 10.1109/ACCT.2014.74.

[41] A. D. Belsare, M. M. Mushrif, M. A. Pangarkar, and N. Meshram, "Classification of breast cancer histopathology images using

texture feature analysis," in *TENCON 2015-2015 IEEE Region 10 Conference*, Nov. 2015, pp. 1–5, doi: 10.1109/TENCON.2015.7372809.

[42] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. K. A. Rao, "Histopathological image classification using discriminative feature-oriented dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 738–751, Mar. 2016, doi: 10.1109/TMI.2015.2493530.

[43] A. Anuranjeeta, K. K. Shukla, A. Tiwari, and S. Sharma, "Classification of histopathological images of breast cancerous and non cancerous cells based on morphological features," *Biomedical and Pharmacology Journal*, vol. 10, no. 1, pp. 353–366, Mar. 2017, doi: 10.13005/bpj/1116.

[44] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Systems with Applications*, vol. 117, pp. 103–111, Mar. 2019, doi: 10.1016/j.eswa.2018.09.049.

[45] Q. Abbas, M. E. Celebi, and I. F. García, "Breast mass segmentation using region-based and edge-based methods in a 4-stage multiscale system," *Biomedical Signal Processing and Control*, vol. 8, no. 2, pp. 204–214, Mar. 2013, doi: 10.1016/j.bspc.2012.08.003.

[46] D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 1, pp. 88–101, Apr. 2014, doi: 10.1016/j.cmpb.2014.01.014.

[47] F. R. Cordeiro, W. P. Santos, and A. G. Silva-Filho, "An adaptive semi-supervised Fuzzy GrowCut algorithm to segment masses of regions of interest of mammographic images," *Applied Soft Computing*, vol. 46, pp. 613–628, Sep. 2016, doi: 10.1016/j.asoc.2015.11.040.

[48] A. Mouelhi, H. Rmili, J. Ben Ali, M. Sayadi, R. Doghri, and K. Mrad, "Fast unsupervised nuclear segmentation and classification scheme for automatic allred cancer scoring in immunohistochemical breast tissue images," *Computer Methods and Programs in Biomedicine*, vol. 165, pp. 37–51, Oct. 2018, doi: 10.1016/j.cmpb.2018.08.005.

[49] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications*, vol. 42, no. 3, pp. 990–1002, Feb. 2015, doi: 10.1016/j.eswa.2014.09.020.

[50] N. D. Khalilabad and H. Hassanpour, "Employing image processing techniques for cancer detection using microarray images," *Computers in Biology and Medicine*, vol. 81, pp. 139–147, Feb. 2017, doi: 10.1016/j.compbiomed.2016.12.012.

[51] S. Kaymak, A. Helwan, and D. Uzun, "Breast cancer image classification using artificial neural networks," *Procedia Computer Science*, vol. 120, pp. 126–131, 2017, doi: 10.1016/j.procs.2017.11.219.

[52] M. Karabatak, "A new classifier for breast cancer detection based on Naïve Bayesian," *Measurement*, vol. 72, pp. 32–36, Aug. 2015, doi: 10.1016/j.measurement.2015.04.028.

[53] F. Wang, S. Zhang, and L. M. Henderson, "Adaptive decision-making of breast cancer mammography screening: A heuristic-based regression model," *Omega*, vol. 76, pp. 70–84, Apr. 2018, doi: 10.1016/j.omega.2017.05.001.

[54] M. R. Mohebian, H. R. Marateb, M. Mansourian, M. A. Mañanas, and F. Mokarian, "A hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning," *Computational and Structural Biotechnology Journal*, vol. 15, pp. 75–85, 2017, doi: 10.1016/j.csbj.2016.11.004.

[55] A. Cruz-Roa *et al.*, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," in *SPIE Proceedings*, Mar. 2014, doi: 10.1117/12.2043872.

[56] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.

[57] 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)*, Jul. 2016, pp. 2560–2567, doi: 10.1109/IJCNN.2016.7727519.

[58] E. Adoui, M, S. A. Mahmoudi, M. A. Larhmam, and M. Benjelloun, "MRI breast tumor segmentation using different encoder and decoder CNN architectures," *Computers*, vol. 8, no. 3, Jun. 2019, doi: 10.3390/computers8030052.

[59] M. U. Dalmış *et al.*, "Using deep learning to segment breast and fibroglandular tissue in MRI volumes," *Medical Physics*, vol. 44, no. 2, pp. 533–546, Feb. 2017, doi: 10.1002/mp.12079.

[60] P. Moeskops *et al.*, "Deep learning for multi-task medical image segmentation in multiple modalities," in *Medical Image Computing and Computer-Assisted Intervention 2016*, Springer International Publishing, 2016, pp. 478–486, doi: 10.1007/978-3-319-46723-8_55.

[61] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.

[62] S. Punitha, A. Amuthan, and K. S. Joseph, "Benign and malignant breast cancer segmentation using optimized region growing technique," *Future Computing and Informatics Journal*, vol. 3, no. 2, pp. 348–358, Dec. 2018, doi: 10.1016/j.fcij.2018.10.005.

[63] J. J. Patel and S. K. Hadia, "An enhancement of mammogram images for breast cancer classification using artificial neural networks," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 2, pp. 332–345, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp332-345.

[64] A. Nomani, Y. Ansari, M. H. Nasirpour, A. Masoumian, E. S. Pour, and A. Valizadeh, "PSOWNNs-CNN: a computational radiology for breast cancer diagnosis improvement based on image processing using machine learning methods," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–17, May 2022, doi: 10.1155/2022/5667264.

[65] G. Serpen and S. Pathical, "Classification in high-dimensional feature spaces: random subsample ensemble," in *2009 International Conference on Machine Learning and Applications*, Dec. 2009, pp. 740–745, doi: 10.1109/ICMLA.2009.26.

[66] S.-H. Min, "A genetic algorithm-based heterogeneous random subspace ensemble model for bankruptcy prediction," *International Journal of Applied Engineering Research*, vol. 11, pp. 2927–2931, 2016

[67] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," *Cancer Informatics*, vol. 2, Jan. 2006, doi: 10.1177/117693510600200030.

[68] M. Hosni, I. Abnane, A. Idri, J. M. Carrillo de Gea, and J. L. Fernández Alemán, "Reviewing ensemble classification methods in breast cancer," *Computer Methods and Programs in Biomedicine*, vol. 177, pp. 89–112, Aug. 2019, doi: 10.1016/j.cmpb.2019.05.019.

[69] S. Saxena and M. Gyanchandani, "Machine learning methods for computer-aided breast cancer diagnosis using histopathology: a narrative review," *Journal of Medical Imaging and Radiation Sciences*, vol. 51, no. 1, pp. 182–193, Mar. 2020, doi: 10.1016/j.jmir.2019.11.001.

[70] E. Sadeghipour, A. Hatam, and F. Hosseinzadeh, "An expert clinical system for diagnosing obstructive sleep apnea with help

from the Xcsr classifier," *Journal of Mathematics and Computer Science*, vol. 14, no. 01, pp. 33–41, Jan. 2015, doi: 10.22436/jmcs.014.01.04.
[71] E. Sadeghipour, A. Hatam, and F. Hosseinzadeh, "Designing an intelligent system for diagnosing diabetes with the help of the Xcsla system," *Journal of Mathematics and Computer Science*, vol. 14, no. 01, pp. 24–32, Jan. 2015, doi: 10.22436/jmcs.014.01.03.
[72] P. Wang *et al*., "Cross-task extreme learning machine for breast cancer image classification with deep convolutional features," *Biomedical Signal Processing and Control*, vol. 57, Mar. 2020, doi: 10.1016/j.bspc.2019.101789.
[73] A. Kumar *et al.*, "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer," *Information Sciences*, vol. 508, pp. 405–421, Jan. 2020, doi: 10.1016/j.ins.2019.08.072.
[74] P. Kaur, G. Singh, and P. Kaur, "Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification," *Informatics in Medicine Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.01.001.
[75] F. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Systems with Applications*, vol. 120, pp. 103–115, Apr. 2019, doi: 10.1016/j.eswa.2018.11.008.
[76] H. Li, S. Zhuang, D. Li, J. Zhao, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomedical Signal Processing and Control*, vol. 51, pp. 347–354, May 2019, doi: 10.1016/j.bspc.2019.02.017.
[77] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep learning to improve breast cancer detection on screening mammography," *Scientific Reports*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-48995-4.
[78] G. Murtaza *et al.*, "Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges," *Artificial Intelligence Review*, vol. 53, no. 3, pp. 1655–1720, Mar. 2020, doi: 10.1007/s10462-019-09716-5.
[79] A. B. Levine, C. Schlosser, J. Grewal, R. Coope, S. J. M. Jones, and S. Yip, "Rise of the machines: advances in deep learning for cancer diagnosis," *Trends in Cancer*, vol. 5, no. 3, pp. 157–169, Mar. 2019, doi: 10.1016/j.trecan.2019.02.002.
[80] N. I. R. Yassin, S. Omran, E. M. F. El Houby, and H. Allam, "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review," *Computer Methods and Programs in Biomedicine*, vol. 156, pp. 25–45, Mar. 2018, doi: 10.1016/j.cmpb.2017.12.012.
[81] M. Saha and C. Chakraborty, "Her2Net: a deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2189–2200, May 2018, doi: 10.1109/TIP.2018.2795742.
[82] Z. Rustam, Y. Amalia, S. Hartini, and G. S. Saragih, "Linear discriminant analysis and support vector machines for classifying breast cancer," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, pp. 253–256, Mar. 2021, doi: 10.11591/ijai.v10.i1.pp253-256.
[83] G. G. N. Geweid and M. A. Abdallah, "A novel approach for breast cancer investigation and recognition using M-Level set-based optimization functions," *IEEE Access*, vol. 7, pp. 136343–136357, 2019, doi: 10.1109/ACCESS.2019.2941990.
[84] R. Guo, G. Lu, B. Qin, and B. Fei, "Ultrasound imaging technologies for breast cancer detection and management: a review," *Ultrasound in Medicine and Biology*, vol. 44, no. 1, pp. 37–70, Jan. 2018, doi: 10.1016/j.ultrasmedbio.2017.09.012.
[85] B. C. Patel and G. R. Sinha, "Mammography feature analysis and mass detection in breast cancer images," in *2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies*, Jan. 2014, pp. 474–478, doi: 10.1109/ICESC.2014.89.
[86] A. K. Singh and B. Gupta, "A novel approach for breast cancer detection and segmentation in a mammogram," *Procedia Computer Science*, vol. 54, pp. 676–682, 2015, doi: 10.1016/j.procs.2015.06.079.
[87] J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 2, pp. 236–251, Mar. 2009, doi: 10.1109/TITB.2008.2009441.
[88] Y. Lee, "Preliminary evaluation of dual-head Compton camera with Si/CZT material for breast cancer detection: Monte Carlo simulation study," *Optik*, vol. 202, Feb. 2020, doi: 10.1016/j.ijleo.2019.163519.
[89] S. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, vol. 125, pp. 1–6, Jul. 2019, doi: 10.1016/j.patrec.2019.03.022.
[90] S. Rahmatinia and B. Fahimi, "Magneto-thermal modeling of biological tissues: a step toward breast cancer detection," *IEEE Transactions on Magnetics*, vol. 53, no. 6, pp. 1–4, Jun. 2017, doi: 10.1109/TMAG.2017.2671780.
[91] Z. Wang *et al.*, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019, doi: 10.1109/ACCESS.2019.2892795.
[92] 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, doi: 10.1109/ACCESS.2019.2904245.
[93] S. Panesar, X. Weng, and S. Neethirajan, "Toward point-of-care diagnostics of breast cancer: development of an optical biosensor using quantum dots," *IEEE Sensors Letters*, vol. 1, no. 4, pp. 1–4, Aug. 2017, doi: 10.1109/LSENS.2017.2727983.

BIOGRAPHIES OF AUTHORS

Ehsan Sadeghi Pour is currently pursuing the Ph.D. degree in Computer Engineering-Artificial Intelligence and Robotics, Islamic Azad University, Kashan branch, Iran. He received Bs in software engineering from the Birjand University of Industries and Mines, Birjand and his M.Sc. in Computer Engineering, Artificial Intelligence from University of Hormozgan, Bandar Abbas in Iran. He is a faculty member at University of Sama in Bandar Abbas, Iran. His research interests include artificial intelligence, machine/deep learning, data mining, computer vision and diagnosis. He can be contacted at email: Ehsan61.ai@gmail.com.

Mahdi Esmaeili received his Ph.D. degree in Computer Engineering at University of Debrecen in Hungary. He received a Bs in software engineering from the University of Isfahan, and his MS in computer science at 1999 in Iran. He is a faculty member of Islamic Azad University (Kashan branch). He has taught in the areas of database and data mining and his research interests include data base, data mining and knowledge discovery, and recent research focusing on the end use of data mining. He has published books about database and information retrieval and also has published articles in international conferences. He can be contacted at email: M.esmaeili@iaukashan.ac.ir.

Morteza Romoozi received his Ph.D. degree in Computer Engineering from Science and research branch of Islamic Azad University in Iran. He received Bs in software engineering from the University of Kashan at 2003, and his MS in computer science at 2006 in Iran. He is a faculty member of Islamic Azad University (Kashan branch). He has taught in the areas of wireless networks, ad hoc and P2P Networks and his research interests include semantic web, information retrieval, and recent research focusing on the mobility model and routing protocol in ad-hoc networks. He has published several articles in international conferences, LNCS series and other springer journals. He can be contacted at email: Mromoozi@gmail.com.