# A comparison of recent progress in breast cancer diagnosis models using machine learning algorithms

**Zhiyu Zhang**

University of Electronic Science and Technology of China, Chengdu, Sichuan, 611700, China

2022080905029@std.uestc.edu.cn

**Abstract.** Breast cancer is the leading cause of mortality among women suffering from cancer, so the accurate diagnosis is important. This review aims to provide a thorough examination of advancements and trends in breast cancer diagnosis by analysing recognized papers published between 2020 and 2023.The paper firstly gives a brief overview of breast cancer, machine learning algorithms, followed by an introduction of basic process for ML in breast cancer diagnosis. After that, by focusing on two emerging trends, hybridization and newly invented modalities, the review introduces existing achievements in the field. Subsequently, it highlights nine notable or novel designs in breast cancer diagnosis, while presenting their comparative properties in a tabular format. Hopefully, this review can equip researchers with valuable insights for future studies and references, helping them gain a better understanding of the field and facilitating further improvements in breast cancer detection and classification.

**Keywords:** Breast Cancer, Computer-Aided Diagnosis, Machine Learning Algorithms.

## 1. Introduction

Over 2.3 million women die annually from breast cancer, which is one of the biggest causes of cancer in adults. In 95 percent of countries, breast cancer is the number one or two leading causes of women's cancer [1]. Breast cancer has a profound impact on individuals, families, and society. It has negative physical and emotional impacts, and it imposes burden financially, significantly decreasing patients' quality of life. The goal of early detection is to bring the phase at which a patient is diagnosed forward, which may increase the chances of survival and cure, and allow easier and more cost-effective treatment [2].

The medical image test is the best way to diagnose the breast carcinoma. A variety of medical imaging techniques are employed in the diagnostic process, encompassing digital mammography (DM), histological images viewed through a microscope, infrared thermography (IRT), magnetic resonance imaging (MRI), and ultrasound (US). Additionally, gene expression datasets, non-image clinical information are gradually being taken into consideration. Machine learning (ML) is a kind of algorithm that the methods and statistics that are used in a computer system to carry out certain tasks. Learning algorithms are utilized in numerous applications that one can encounter on a daily bases.ML algorithms find application in diverse domains. The main benefit of using machine learning is that the algorithm can perform the job automatically when it knows how to deal with the data, which can be employed to aid medical practitioners in breast cancer diagnosis. According to the Scopus database, the documents

dedicated for machine learning techniques for breast cancer diagnosis between 2010 to 2019 had been climbing up by years, and there sprung out various kinds of ingenious and efficient techniques. Brief introductions of most commonly used algorithms are as follows.

Support Vector Machines (SVM) is one of the most widely applied methods for classification and regression tasks. It is a supervised learning model that analyzes data and finds the optimal hyperplane or decision boundary to separate different classes. In general, its mission is balancing the trade-off between maximizing the margin and minimizing the misclassification of data points when using appropriate kernel functions to transform the input vectors.

In recent decades, several enhancements to the SVM algorithm have emerged, including Quantum Support Vector Machine, least square support vector machine, twin SVM and Lagrangian SVM, etc. [3]. The maximum accuracy attained using SVM can reach 100% [4]. Being a powerful classifier, efficient processor of high-dimension features and small sample data and strong mathematical model, SVM has been widely applied in natural language process (NLP), image recognition, medical field etc.

Artificial Neural Network (ANN) is a modality which imitates architectures and functionality of the biological neural system. With A vast network of correlated computational neurons, it transmits and process information through weights. The working principle of ANN involves calculating forwarding input data through multiple layers using a predefined set of weights and activation functions. Then, the backpropagation algorithm compares the output with the desired output, adjusts the weights to minimize the default loss function. This process iterates until a specified level of accuracy or convergence is achieved.

CNN is a structure based on ANNs that has been improved and optimized from traditional fully connected neural networks. It employs specific components such as convolutional layers and pooling layers to process data with spatial structures, such as images and audio. In CNN, the convolutional layers extract features from input data using convolutional kernels (filters). These layers reduce the parameter count of the network through parameter sharing and local perception, while preserving the spatial information of the data. On the other hand, the pooling layer is employed to reduce the size of the feature map and to get more prominent characteristics. Notably, Convolutional Neural Networks (CNNs) stacked by spiking number of layers have shown exceptional performance in multiple domains. For instance, in the field of object detection, algorithms like Faster R-CNN and YOLO utilize CNNs to in image detection can provide their positions and categories. These algorithms have broad applications in autonomous driving, security surveillance, and other domains.

K-Nearest Neighbours (KNN) is a fundamental machine learning algorithm commonly used for many classical machine learning problems. Its working principle is quite simple and intuitive: given a new data point, the KNN algorithm finds the K nearest neighbours to that point in the training dataset and predicts or decides based on the labels of those neighbours. In Raghavendra et al. 's model [5], the KNN is a classifier that obtained a maximum average precision of 98.69%. And it's the third frequently used technique in CAD of breast cancer.

Decision Tree (DT) is one of the most common methods in the field of machine learning when dealing with the basic issues. It makes decisions based on a tree-like structure, where a series of checks and branches are applied to input data in order to achieve prediction or decision-making. In its procedure, the algorithm selects features based on information gain rate, Gini index or etc. Then it recursively builds the following subbranches as the decision model, during which time it adopts certain kind of pruning technique to avoid overfitting. DTs are easily understandable, and they can deal with both discrete and sequential representations, as well as dealing with missing data. Moreover, decision trees offer an effective approach to feature selection, allowing for the identification of the most impactful features that influence the target variable. Decision trees are widely applied in various fields including medical diagnosis, financial risk assessment, and marketing. They are powerful tools that help us understand and tackle complex decision problems. In Nguyen et al. [6] adopted a model using RF, which aggregates the decision of each independent tree by voting, attained classification accuracy of 100% and approximately 99.8% respectively in the best and average run. And Figure 1 shows the basic process for ML in breast cancer diagnosis.
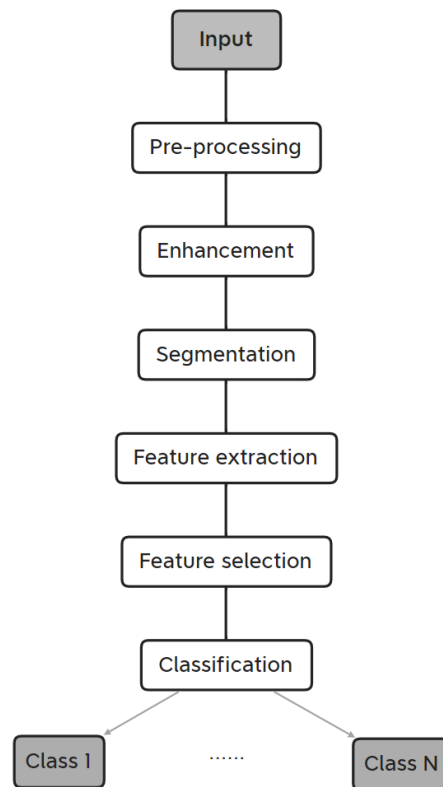
**Figure 1.** Basic process for ML in breast cancer diagnosis.

Preprocessing in machine learning refers to the process of transforming, cleaning, and preparing raw data before applying machine learning algorithms. It includes steps such as data cleaning to handle noise, missing values, and outliers, feature scaling to bring features to a consistent scale, feature selection to reduce dimensionality, feature transformation for extracting valuable information, data splitting into training and testing sets, and label encoding for converting categorical variables into numerical representations. Enhancement in machine learning refers to techniques and approaches aimed at improving the performance, accuracy, or capabilities of a machine learning model. Some common enhancement techniques include feature engineering, hyperparameter tuning, ensemble methods, regularization, optimization, data augmentation, and transfer learning. Segmentation in machine learning involves dividing an input data or image into distinct segments or regions based on specific criteria or features. It is commonly used in computer vision tasks to identify and delineate objects or regions of interest. Feature extraction in machine learning involves transforming raw or high-dimensional data into a more compact and meaningful representation by selecting or extracting relevant features. This process reduces dimensionality, removes noise, and improves model performance and interpretability. Techniques like PCA, LDA, manifold learning, deep feature extraction, and domain-specific methods are commonly used for feature extraction. Feature selection in machine learning consists of choosing a subset of corelated and informative features from a larger set. It is targeted at simplify the algorithm, improve computational efficiency, reduce overfitting, and enhance interpretability. Classification in machine learning are algorithms or models that categorize input data into different classes. Some depend on models that are supported by mathematics as the underlying logic, such as SVM and Discriminative analysis (DA). Some are relatively more intuitive like Decision Tree or K nearest neighbor algorithms. Some simulate biological neural systems with remarkable performance for unknown reasons like ANN.

## 2. Method

### 2.1. Recent trend

Recent year, Hybridization and fusion has become a preferable trend. Moreover, many novel and ingenious network structures have been proposed. Yao et al. [7] introduced a novel deep learning architecture that combined two distinct neural networks, DenseNet (NN that accesses posture of the human body), and LSTM (long short-term memory). Their model incorporated a specialized perceptron attention mechanism commonly employed in Natural Language Processing (NLP) applications, which effectively integrated the image features extracted from both networks. Wu, E. et al. [8] employed a deep learning method of generative adversarial networks (GANs) that were introduced to synthesize authentic images as input datasets, to solve the problem of limited datasets and annotations.

### 2.2. Trails and objective of this paper

Two trails that this paper will be following are the improvements in the field of machine learning where researchers have achieved fruitful accomplishments currently, and the instructive ideas that have been recently brought up. The first trail will lead to three categories of Multimodal, Hybridization and Ensemble learning and phased optimization, while the second trail will lead to new model application. This paper aims to collecting and analyzing popular papers that published between 2020-2023 to give a better overview in the domain of breast cancer diagnosis, hopefully better preparing researchers for future study or reference.

## 3. Results and discussion

**Table 1.** Review of well-recognized or novel papers published from 2021 to 2023.

| References | Model type | Datasets | Outputs | Numerical results | Advantages | Limitations |
|---|---|---|---|---|---|---|
| [9] | CNN (MultiModal fusion learning) | Related 18 clinical information combined with corresponding 17,046 processed 2D MIP breast images | probability of malignancy | (Five-run ensemble of Probability Fusion) -AUC of 0.888 -Specificity of 95% -Sensitivity of 51.3 | Leveraged the clinical risk factor data. | -datasets were private, derived from one institution. - hyperparameter tuning and architecture design choices need further exploring. |
| [10] | CNN (hybridization) | -mini-DDSM dataset -BUSI dataset -BUS2 dataset | abnormality and malignancy classification | -mini-DDSM dataset: accuracy 98.00 sensitivity 98.00 specificity 98.00 | - Hybridization exploit advantages of each network -the valve-like control reduces | hybridization with best performance was more time-consuming than AlexNet, ResNet, |

| | | | | | |
|---|---|---|---|---|---|
| | | | -BUSI dataset: accuracy 93.70 sensitivity 94.62 specificity 92.74 -BUS2 dataset 98.00 accuracy 97.00 sensitivity 99.00 specificity | misclassifications | MobileNet, ShuffleNet independently |
| [11] | CNN (Ensemble learning) | -Synthetic Gene Image -Breast histopathology image Dataset | Possibility of healthy and tumor | -Accuracy 98.08% -F1-Score 0.988% -Recall 0.9920% -Precision 0.9841% -Sensitivity0.9920% -pecificity0.9355% | performance of the training using EWT and VMD decomposed modes outperformed original CNN | hardware implementation is required to validate its superior performance. |
| [12] | CNN (ResNet18+ ELM) | -CBIS-DDSM -MIAS -INbreast datasets | Three-class classification | (Overall classification accuracy) -MIAS 98.137% -DDSM 97.193% -INbreast 98.266% | -exploit ELM's property of speediness and efficiency -optimize the ELM with crow search algorithm. | Overall metrics other than accuracy are not specified in number |
| [13] | CNN+ transfer learning | MIAS datasets | Three-class classification | (Overall, in 80-20 method) Accuracy 98.96% Sensitivity 97.83% Specificity 99.13% Precision 97.35% F-score 97.66% | -enhancement of classification accuracy -speeding up of the training process | lack of comparison with related work |

| | | | AUC 0.995 | | |
|---|---|---|---|---|---|
| [14] | IMPA-ResNet50 | -5283 images in CBIS-DDSM -1290 in MIAS datasets | Binary (benign or cancer) | -CBIS-DDSM dataset: Accuracy 98.32% Sensitivity 98.56% Specificity98.68% -MIAS dataset: Accuracy 98.88% Sensitivity 97.61% Specificity98.40% | -propose an optimization searching for the hyperparameters for the network which outperforms state-of-the-art optimization | -comparatively computationally expensive -time-consuming -the datasets tested are limited - Only succeeded in specifying ResNet50 hyperparameters |
| [15] | AnoGAN + DensNet | BreaKHis dataset | Binary (benign or malignant) | (The best run) Accuracy 99.13% F1score of 99.38% | achieved satisfactory classification performance for coarse-grained high-resolution images | the preprocessing of the input images neglects the data imbalance that are common in medical domain - hyperparameters, and the size of patches etc. Need further tuning |
| [16] | decision tree-based ensemble learning | Wisconsin breast cancer database | Binary (benign or malignant) | (For both Random Forest and Extra Trees algorithms) Accuracy 100% Sensitivity 100% Specificity 100% | -high accuracy -strong interpretability, user-friendly and simple to understand | feature-based learning technique requiring tabular data that are scarce. |
| [17] | Improved CNN+GCN | 322 mammographic images from | Binary (benign or malignant) | Sensitivity of 96.20±2.90%, Specificity of 96.00±2.31% and Accuracy | Each image is described by both its image-level and its neighbour | -not reliably to interrogate heterogeneous data |

| miniMIAS dataset | of 96.10±1.60% | characteristics, achieving high accuracy | -the employed datasets were not large |
|---|---|---|---|

Various diagnosis models will be interpreted as follows and their brief summaries are represented in Table 1, depending upon the Model type, Datasets, Outputs, Numerical results, Advantages and Limitations.

Gregory Holste et al. [9] utilized an end-to-end MultiModal Fusion method to improve the results of classification of breast cancer detection, integrating both MRI images and non-image information including visual data, larithmic and clinical data cancer risk factors. They devised three intuitive ways to fuse two modalities with one fusing two outcome number, one fusing learned features of images directly with non-image data and one fusing both learned features of two models. Among three methods, the Learned Feature Fusion, which fused intermediate learned features from each modality, outperformed the others. In addition, they found that different fusion operations also effected the performances. The limitations are that their data were collected from one institution, the deficiency in non-image data and the identical hyperparameters between three ways of fusion.

Adyasha Sahu et al. [10] contributed 5 different hybrid classifiers for more accurate breast cancer detection. In their overall architecture, they designed two classification modules with one only activated when a probability-based weight factor surpassed a threshold, so as to avoid misclassification. And the five hybrid frameworks are based on the different matches for the two classifiers. For classifiers 1 and classifiers 2, they were respectively: MobileNetV2-ResNet18, VGG16 -ResNet18, ResNet18-AlexNet, ShuffleNet-AlexNet and ShuffleNet-ResNet18. They found out that the approaches devised perform better than other approaches in all of the performance metrics both in abnormality detection and malignancy detection. Their limitation is that the hybridization with best performance was more time-consuming than AlexNet, ResNet, MobileNet, ShuffleNet independently.

After the preprocessing of the mammograms, Chakravarthy et al. [11] model extracted features with Resnet18.Subsequently,they fed these features into an optimized classifier, Extreme Learning Machine(ELM) with an improved Crow-Search Optimization Algorithm (CSOA), which was devised by a Single-Hidden-Layer Feedforward Networks(SLFNs)that was improved by a controlled operator and chaotic-maps generated randomness, so as to make use of ELM's property of speediness and efficiency in solving non-linear classification tasks while improving its drawbacks of reduced classification accuracy due to the stochastic initialization of input weight and hidden bias. Compared with current models that used the same datasets, the results showed that ICS-ELM with sine chaotic map had the highest overall accuracy, lowest total misclassification cost and highest kappa compared with other combinations of its disassembled parts and ICS-ELM with logistic chaotic map. And The performance of the training using EWT and VMD decomposed modes exceeded convolutional neural networks trained in the most basic way.

Abhishek Das et al. [12]Used 1-D breast histopathology dataset and Synthetic Gene Dataset from Kaggle as input .In order to utilize both the gene expression data and the effectiveness of CNN ,they inverted the 1-D datasets into 2-D images with approaches of tSNE (t-Distributed Stochastic Neighbor Embedding), Convex Hull and bounding rectangle .After obtaining the 2-D images datasets, they designed three CNN pipelines: in the first pipeline, they directly fed the 2-D images datasets obtained into the CNN; in the last two pipelines, they respectively employed Empirical wavelet transform (EWT) and Variational Mode decomposition (VMD) before feeding datasets into CNNs .Following this phase, they concatenated the predictions of these three classifiers and fed the outcomes into a two-layered MLP with two output neurons for IDC(images containing cancer cells ) and Non-IDC classification. They concluded that their ensemble architecture performed better and the independent base classifiers, and surpassed other state-of-the-art method. The performance of the training using EWT and VMD decomposed modes exceeded original CNNs. The proposed method demonstrates superior performance

if carried out under the emulation circumstance. Their limitation lie in that the hardware implementation is required to validate its limitations.

Saber et al [13] proposed a series of pre-processing techniques (noise elimination, histograms balance, morphology analysis, segmentation, scaling, data division, and enhancement) to enhance the quality of raw images. Subsequently, the processed pictures were fed to CNN (respectively Inception-V3, VGG19, VGG16, ResNet50, InceptionV2-ResNet), which were pre-pretrained on ImageNet. It is found out that the transfer learning of the VGG16 model can be used to detect breast cancer effectively. Houssein et al. [14] used a pretrained ResNet50 hybridized with an improved marine predator algorithm (IMPA) using opposition-based learning strategy, which targeted at finding the optimal hyperparameters for the neural networks by three stages simulated the chasing movement in a predation. The proposed model reached high accuracy, sensitivity and specificity on both the CBIS-DDSM dataset and the MIAS dataset, outperforming other state-of-the-art approaches and other optimization algorithms.

Based on patches for high resolution pathological images, a new approach has been used. GANs have been proposed to overcome a major issue in patch-based image classification. In order to recognize the most distinguishing pathological spots of breast carcinoma from malignant images, MAN et al. [15] acquired a GAN for a good picture patch with a GAN, so that there is a significant difference when studying the distribution of a harmful picture patch with this GAN. Then, they fed the corrected patches into DenseNet121 to fulfill the task of classifying the target. The findings indicate that AnoGAN provides a further improvement in the precision of the approach to screening discriminatory patches, with respect to both individual level precision and visualization level precision, as compared with a approach that does not employ AnoGAN for patches screening.

Ghiasi et al. [16] A number of RF/ET categorization models have been studied in combination with GINI-based CARTs in the Model Structure. In their thesis, RF has been used as an ensemble model, while the Classification and Regression Tree (CART) has been used as a model. The results show that RF models have 4 to 10 CARTs and 3 to 9 ET models. CARTs are able to accurately, accurately, accurately and accurately for WBCD types in all cases, as demonstrated by the results, which achieve a high precision and are easily understandable by the use of tree decision-making models.

In Zhang et al. [17]'s model, they added batch normalization and dropout layers to a basic convolutional network, replaced the maxpooling layers in CNN with rank-based stochastic pooling (RSP) layers, and used "or" operator to combine the outcomes of this improved CNN with the outcomes of another pipeline of GCN to exploit the relationship representations in the images. Specifically, the GCN generated an adjacency matrix (ADM) by implementing a cosine similarity (CS)-based kNN. The ADM were passed to a 2L-GCN to receive a linear projection to get the final outcomes of GCN. Jointly, the model attained three mainstream numerical results all around 96%.

## 4. Comparison

### 4.1. Analysis of pros and cons

The listed models are typical examples of the following 4 models: the first paper mentioned in 3. [9] is the example for Multimodal, and the second and the third [10,11] for Hybridization and Ensemble learning, the fourth to the sixth [12-14] for phased optimization, and the rest [15-17] for new model application. The last category is presented here for idea inspiration and the rest three categories show recent progress made in the fledging field of breast cancer diagnosis.

Advantages for model using multimodal approaches include comprehensive information, more accurate results, better robustness, and improved user experience. However, there are also disadvantages such as higher data collection and processing costs, increased algorithm complexity, and the need for ample data and computational resources. In the first paper mentioned in 3. [9], high AUC and Specificity and complex and large architecture are respectively its pros and cons.

Advantages for Hybridization and Ensemble learning include high adaptation to more versatile features and low risk of overfitting. However, their increased difficulty in combination, training and optimization needs to be taken into consideration. Both the second paper [10] and the third [11] achieved

marvelously high numerical results and made full use of the advantages of every single model or algorithms. On the other hand, both models adopted the transfer learning technique to avoid the time-consuming training and optimization time, which may neglect the potential data distribution difference according to different tasks. Phased optimization can cover all stages mentioned in 1. 2.. Usually, models belong to this category borrow the classical model as the overall architecture and add regional modification to improve the performance of the targeted stage. Although this kind of improvement can achieve spectacular performance (like what cross validation does), it's contribution to the total structure is not likely to be ground-breaking. Additionally, because the algorithms that can be selected to improve the existing model is vast without certain limitation, it requires a large number of experiments in various models or clever inspiration. The fourth paper [12] added an improved Crow Search algorithms to an Extreme Learning Machine which was not commonly used for classification in CNN, the fifth paper [13] deployed transfer learning to optimize the performance of each network, while the sixth paper [14] aimed to optimize the hyperparameters with an improved Marine Predators Algorithm. By introducing novel algorithms into the model and making appropriate adjustments to the algorithms to better fit the circumstances, these three models all performed well in terms of nearly all metrics. The drawbacks vary from time-consuming to computational expensive depending on the additional algorithm the model introduced.

The last category covers creative model designs that are rarely adopted for breast cancer diagnosis tasks or models that incorporates newly invented architecture. Innovation has the power to break free from existing constraints and breathe new life into existing domains, which cannot be ignored by researchers. However, new domains also mean there can be a long and twisted road waiting to stumble the travelers. The eighth paper [16] is the first time that a simple visualization of an ET approach combined with a Classification and Regression Tree (CART) for WBCD datasets. The seventh paper [15] and the ninth paper [17] that respectively employed GANs and GCN for segmentation and a parallel pipeline are the representation of models that incorporates newly invented architecture. According to their experiments, they show a potential of a very promising future. Whereas the new models still lack large number of tests on different tasks and on different datasets to ensure their performance.

*4.2. Advice on determining models for clinical application*
As a general, researchers or relevant practitioners who develop the computer-aided system for breast cancer diagnosis should consider two main factors: resource and ranked target outcomes.

Understandably, the handy resource constrains our model choice. Firstly, with data being an important resource in the information age, choices of models are influenced by the type of datasets if the data form is decided. The model designed to deal with image data like mammograms has discrepancy between model designed to deal with gene expression data or non-image clinical information. On the other hand, if the data form is not decided, CAD developers should leverage the pros and cons of the specific screening techniques that they utilize to acquire training data and to perform on patients. For example, datasets in form of Mammograms have massive amount of data for comparison and learning while it put the patients in the risk of radiation, and it is not ideal for young patients. And for those who adopt gene or RNA data as the input, they take the advantage of cancer's cellular variations, while suffer from high dimensionality in feature selection algorithms. Anyway, the developers should try to conjure up a corresponding algorithm to implement on certain types of input they choose with constraint resource. Additionally, while most CAD software are installed on computers, the prevalence in Embedded system urges the CAD developers to make application of CAD more portable and less space consuming. As for storage space control, deep learning networks need to be carefully considered for they usually occupy a large deposit. Of course, even in computers, a smaller installing requirement is always preferable. Other resource constraints, including project funds for the hardware and human resource, are also the factors need being taken into account when choosing algorithms for clinical application. And when the CAD developers evaluate the model, out of practical needs in clinical application, we not only have to balance between pros and cons of data acquisition as well as time and space, but also need to design the metrics of the model according to the target outcomes ranked by

priority value. For instance, precision and recall may have a trade-off relationship in certain scenarios, where improving one metric might result in a decrease in the other metric. Under this circumstance, the choice of evaluation of the model depends on which metrics rank high in your clinical diagnosis. Furthermore, it depends on whether the developers consider misdiagnosing patients as positive as a serious issue or overlooking positive patients as a serious issue.

## 5. Conclusion

The objective of this paper is to gather and analyze widely recognized papers published from 2021 to 2023, aiming to provide a representative overview of advancements and trends in the domain of breast cancer diagnosis using ML, thus to equip researchers with valuable insights for future studies and references. By tracing two trails of trend, which can be further developed into four categories, this paper firstly introduced the existing achievements and then elucidated 9 notable or novel designs in the field of breast cancer diagnosis while giving comparable properties of each model in tabular form. Successively, the paper gave the analytical comparison of the mentioned literatures.

Although it is relatively easy for us to examine and judge each stage of preprocessing, segmentation, classification etc. independently, the integrated models are more than a simple arrangement and combination of various processing stages. To design a model that somehow makes improvement in specific aspects need ingenious invention or association and numerous trial and error. And their applications also vary depending on the different applicability of these models.

## References

[1]  Anon, n.d. Who launches new roadmap on breast cancer. [online] World Health Organization. Available from: https://www.who.int/news/item/03-02-2023-who-launches-new-roadmap-on-breast-cancer [Accessed 31 Aug. 2023].

[2]  Anderson, B.O.,, Braun, S.,, Lim, S.,, Smith, R.A.,, Taplin, S., and Thomas, D.B., 2003. Early detection of breast cancer in countries with limited resources. The Breast Journal, 9(s2).

[3]  Chandra, M.A., and Bedi, S.S., 2018. Survey on SVM and their application in Image Classification. International Journal of Information Technology, 13(5), pp.1–11.

[4]  Jian, W.,, Sun, X., and Luo, S., 2012. Computer-aided diagnosis of breast microcalcifications based on dual-tree complex wavelet transform. BioMedical Engineering OnLine, 11(1).

[5]  Raghavendra, U.,, Rajendra Acharya, U.,, Fujita, H.,, Gudigar, A.,, Tan, J.H., and Chokkadi, S., 2016. Application of Gabor wavelet and locality sensitive discriminant analysis for automated identification of breast cancer using digitized mammogram images. Applied Soft Computing, 46, pp.151–161.

[6]  Nguyen, C.,, Wang, Y., and Nguyen, H.N., 2013. Random Forest classifier combined with feature selection for breast cancer diagnosis and prognostic. Journal of Biomedical Science and Engineering, 06(05), pp.551–560.

[7]  Yao, H.,, Zhang, X.,, Zhou, X., and Liu, S., 2019. Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification. Cancers, 11(12), p.1901.

[8]  Wu, E.,, Wu, K.,, Cox, D., and Lotter, W., 2018. Conditional infilling gans for data augmentation in mammogram classification. Image Analysis for Moving Organ, Breast, and Thoracic Images, pp.98–106.

[9]  Holste, G.,, Partridge, S.C.,, Rahbar, H.,, Biswas, D.,, Lee, C.I., and Alessio, A.M., 2021. End-to-end learning of fused image and non-image features for improved breast cancer classification from MRI. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW).

[10]  Sahu, A.,, Das, P.K., and Meher, S., 2023. High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets. Biomedical Signal Processing and Control, 80, p.104292.

[11]  Abdollahi, J.,, Davari, N.,, Panahi, Y., and Gardaneh, M., 2022. Detection of metastatic breast cancer from whole-slide pathology images using an ensemble deep-learning method. Archives of Breast Cancer, pp.364–376.

[12]  Sannasi Chakravarthy, S.R., and Rajaguru, H., 2022. Automatic detection and classification of mammograms using improved extreme learning machine with Deep Learning. IRBM, 43(1), pp.49–61.

[13]  Saber, A.,, Sakr, M.,, Abo-Seida, O.M.,, Keshk, A., and Chen, H., 2021. A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. IEEE Access, 9, pp.71194–71209.

[14]  Houssein, E.H.,, Emam, M.M., and Ali, A.A., 2022. An optimized deep learning architecture for breast cancer diagnosis based on improved Marine Predators algorithm. Neural Computing and Applications, 34(20), pp.18015–18033.

[15]  Man, R.,, Yang, P., and Xu, B., 2020. Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. IEEE Access, 8, pp.155362–155377.

[16]  Ghiasi, M.M., and Zendehboudi, S., 2021. Application of decision tree-based ensemble learning in the classification of breast cancer. Computers in Biology and Medicine, 128, p.104089.

[17]  Zhang, Y.-D.,, Satapathy, S.C.,, Guttery, D.S.,, Górriz, J.M., and Wang, S.-H., 2021. Improved breast cancer classification through combining graph convolutional network and convolutional neural network. Information Processing &amp;amp; Management, 58(2), p.102439.